

Comparison of SST diurnal variation models over the Tropical Warm Pool region

Article

Published Version

Zhang, H., Beggs, H., Merchant, C. J., Wang, X. H., Majewski, L., Kiss, A. E., Rodriguez, J., Thorpe, L., Gentemann, C. and Brunke, M. (2018) Comparison of SST diurnal variation models over the Tropical Warm Pool region. *Journal of Geophysical Research: Oceans*, 123 (5). pp. 3467-3488. ISSN 2169-9291 doi: <https://doi.org/10.1029/2017JC013517> Available at <https://centaur.reading.ac.uk/76724/>

It is advisable to refer to the publisher's version if you intend to cite from the work. See [Guidance on citing](#).

To link to this article DOI: <http://dx.doi.org/10.1029/2017JC013517>

Publisher: American Geophysical Union

All outputs in CentAUR are protected by Intellectual Property Rights law, including copyright law. Copyright and IPR is retained by the creators or other copyright holders. Terms and conditions for use of this material are defined in the [End User Agreement](#).

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

RESEARCH ARTICLE

10.1029/2017JC013517

Comparison of SST Diurnal Variation Models Over the Tropical Warm Pool Region

Key Points:

- Four SST diurnal variation models have been evaluated and intercompared, against v3 MTSAT-1R SST measurements
- The sensitivity of MTSAT-1R data to changes in drifting buoy SSTs is tested; scaled MTSAT-1R data are also used in the model comparison
- ZB05 may best capture the overall average dSST cycles and spatial distribution under most DV and wind conditions

Correspondence to:

H. Zhang,
haifeng.zhang@ymail.com

Citation:








Zhang, H., Beggs, H., Merchant, C. J., Wang, X. H., Majewski, L., Kiss, A. E., et al. (2018). Comparison of SST diurnal variation models over the Tropical Warm Pool region. *Journal of Geophysical Research: Oceans*, 123, 3467–3488. <https://doi.org/10.1029/2017JC013517>

Received 29 SEP 2017

Accepted 17 APR 2018

Accepted article online 23 APR 2018

Published online 17 MAY 2018

Haifeng Zhang^{1,2,3} , **Helen Beggs**⁴, **Christopher J. Merchant**⁵ , **Xiao Hua Wang**^{1,2} , **Leon Majewski**⁴, **Andrew E. Kiss**^{3,6} , **José Rodríguez**⁷ , **Livia Thorpe**⁷, **Chelle Gentemann**⁸ , and **Michael Brunke**⁹ 

¹The Sino-Australian Research Centre for Coastal Management, University of New South Wales, Canberra, ACT, Australia,

²School of Physical, Environmental and Mathematical Sciences, University of New South Wales, Canberra, ACT, Australia,

³ARC Centre of Excellence for Climate System Science, The University of New South Wales, Sydney, Australia, ⁴Bureau of Meteorology, Melbourne, Vic, Australia, ⁵Department of Meteorology, University of Reading, Reading, UK, ⁶Research School of Earth Sciences, Australian National University, Canberra, ACT, Australia, ⁷Met Office, Exeter, UK, ⁸Earth and Space Research, Seattle, WA, USA, ⁹Department of Hydrology and Atmospheric Sciences, University of Arizona, Tucson, AZ, USA

Abstract Four sea surface temperature (SST) diurnal variation (DV) models have been compared against Multifunctional Transport Satellite-1R (MTSAT-1R) SST measurements over the Tropical Warm Pool (TWP) region (90°E–170°E, 25°S–15°N) for 4 months from January to April 2010. The four models include one empirical model formulated by Chelle Gentemann (hereafter CG03), one physical model proposed by Zeng and Beljaars in 2005 (ZB05) and its updated version (ZB + T), and one air-sea coupled model (the Met Office Unified Model Global Coupled configuration 2, GC2) with ZB05 warm layer scheme added on top of the standard configuration. The sensitivity of the v3 MTSAT-1R data to the “true” changes in SST is first investigated using drifting buoys and is estimated to be 0.60 ± 0.05 . This being significantly different from 1, the models are validated against MTSAT-1R data and the same data scaled by the inverse of the sensitivity (representing an estimate of the true variability). Results indicate that all models are able to capture the general DV patterns but with differing accuracies and features. Specifically, CG03 and ZB + T underestimate strong (> 2 K) DV events’ amplitudes especially if we assume that sensitivity-scaled MTSAT-1R variability is most realistic. ZB05 can effectively capture the DV cycles under most DV and wind conditions, as well as the DV spatial distribution. GC2 tends to overestimate small-moderate (< 2 K) DV events but can reasonably predict large DV events. One to three hour lags in warming start and peak times are found in GC2.

1. Introduction

The importance of incorporating sea surface temperature (SST) diurnal variations (DVs) into numerical weather prediction (NWP) or climate models is becoming more recognized (e.g., Masson et al., 2012; Stuart-Menteth et al., 2003; Takaya et al., 2010). Including SST DV effects has been demonstrated to more accurately represent the air-sea interaction and therefore is expected to enhance air-sea coupled, NWP, and climate model performance (e.g., Brunke et al., 2008; Clayson & Bogdanoff, 2013; Halpern & Reed, 1976; Li et al., 2001; Marullo et al., 2016). For instance, Clayson and Bogdanoff (2013) have shown that including SST DV effects results in up to 10 W m^{-2} yearly average heat flux difference over significant portions of the tropical oceans. Also, according to Brunke et al. (2008), after introducing a SST DV scheme, the Community Atmosphere Model (CAM3.1) is able to better predict the diurnal cycle in air temperature and precipitation. Some studies have demonstrated that incorporating the SST DV signal can improve a climate model on long time scales. For instance, Masson et al. (2012) showed that in their global coupled ocean-atmosphere general circulation model called SINTEX-F2 (the Scale Interaction Experiment-Frontier version 2), the prediction of ENSO amplitude, frequency, and skewness is significantly better if SST DV effects are taken into account. However, unfortunately, DV effects are still missing in many coupled, NWP or climate models.

There are several numerical SST DV models available. They range in complexity from those based on empirical relationships (e.g., Gentemann et al., 2003; Price et al., 1987; Webster et al., 1996), to single layer or multi-layer bulk models (e.g., Fairall et al., 1996; Price et al., 1986; Zeng & Beljaars, 2005), and to turbulent closure models (e.g., Kantha & Clayson, 1994; Karagali et al., 2017). Kawai and Wada (2007) provided a systematic

review of many of the above models along with several others and suggested that most models are able to resolve the general DV patterns but with differing accuracies and other issues. Other DV model types also include, but are not restricted to, transient models (e.g., Soloviev & Lukas, 2013), air-sea coupled models (e.g., Noh et al., 2011), and physical-empirical hybrid models (e.g., Gentemann et al., 2009; Weihs & Bourassa, 2014).

This paper compares four DV models against version 3 Multifunctional Transport Satellite-1R (MTSAT-1R) SST measurements (produced by the Australian Bureau of Meteorology, hereafter the Bureau) over the Tropical Warm Pool (TWP) region (defined in this study as 90°E–170°E, 25°S–15°N) for 4 months from January to April 2010. The TWP study domain, which is the ocean water mass located in the western Pacific Ocean and eastern Indian Ocean, is considered an excellent area to conduct DV studies because of its globally highest annual average SST, relatively calm winds, high cloud-free values of solar shortwave insolation (SSI), and frequent large amplitude DV events (Zhang et al., 2016a). Although there are many DV models available, and they have all been validated against in situ or satellite observations when proposed or in later papers, few publications conduct intercomparison between different models (e.g., Bellenger & Duvel, 2009; Karagali & Høyer, 2013) and none specifically over the TWP region. If NWP or climate modelers consider incorporating DV effects, selecting an appropriate DV scheme will be a key preliminary step. Therefore, the authors believe that an intercomparison between different DV models could potentially provide useful information and be of interest to NWP or climate modelers. Four models being investigated in this study are one empirical model based on Gentemann et al. (2003; hereafter CG03), two physical models from Zeng and Beljaars (2005; hereafter ZB05) and Takaya et al. (2010; hereafter ZB + T), and one air-sea coupled model (the Met Office Unified Model Global Coupled configuration 2, hereafter GC2; Williams et al., 2015) with Artale et al. (2002) cool skin and the ZB05 warm layer schemes added on top of the standard configuration. The cool skin is a layer in the upper few hundredths of millimetres of the ocean caused by the combined cooling effects of the net longwave radiation, the sensible heat flux, and the latent heat. The cool skin is of the order of 0.1–0.5 K and is almost always present, although its total effect may be compensated by the presence of a warm layer. Warm layers occur during the day when temperature stratification caused by absorption of the solar flux is sufficiently strong to suppress shear-induced mixing and can be of the order of several K (Fairall et al., 1996).

This DV model comparison work is part of the Group for High Resolution Sea Surface Temperature (GHRSSST) Tropical Warm Pool Diurnal Variability (TWP+) project (Beggs et al., 2012). The TWP+ project was launched in 2008 by the GHRSSST Diurnal Variability Working Group with the aim to quantify SST DV over the TWP region and systematically compare different DV models against satellite measurements over the time span of 4 months from January to April 2010. The initial model selection for the TWP+ project, including CG03, ZB05, and ZB + T, was based on their different modeling methods and low computational cost. The GC2 model is added in this study to evaluate the inclusion of warm layer and cool skin schemes to a latest-generation coupled general circulation model.

The structure of the rest of this paper is as follows: section 2 gives details of the models and describes the data sets used; section 3 investigates the sensitivity of MTSAT-1R data to changes in in situ SST measurements; section 4 presents the results of the DV model validation and intercomparison; discussion and conclusions are provided in section 5.

2. Models, Data, and Methods

Before introducing the models, the SST definitions used in this study are summarised. According to Donlon et al. (2007) and the GHRSSST website (<https://www.ghrsst.org/>), the skin SST, SST_{skin} , is the “temperature measured by an infrared radiometer typically operating at wavelengths 3.7–12 μm that represents the temperature within the conductive diffusion dominated sublayer at a depth of ~ 10 –20 μm .” Subskin SST, $SST_{subskin}$, is representative of the SST at the bottom of the SST_{skin} temperature gradient, i.e., the cool skin layer. The foundation SST, SST_{fnd} , is the SST at a depth so that it is free from diurnal variability, which corresponds to the minimum $SST_{subskin}$ during a solar day. More details can be found in Donlon et al. (2002, 2007) and the GHRSSST website.

In this study, we define “dSST” as the hourly $SST_{subskin}$ minus SST_{fnd} and “dSST_{max}” as the maximum dSST during a solar day. Readers should note that two types of SST_{fnd} data, slightly different from the exact Donlon et al. (2007) definition, are used in this study and are described in detail in section 2.3.

2.1. Models

2.1.1. The CG03 Model

An empirical DV model was formulated in Gentemann et al. (2003) using nonlinear least squares regression of passive microwave SST and surface wind speed data, both obtained from the Tropical Rainfall Measuring Mission (TRMM) Microwave Imager (TMI), and top-of-atmosphere modeled insolation data. The model determines the SST DV cycle from daily average insolation and wind speed only; there is no explicit consideration of latent or sensible heat fluxes, although these will presumably be captured to some extent by the empirical dependence on wind speed.

One limitation of the original passive microwave SST-derived Gentemann et al. (2003) formula was that the TRMM satellite's equatorial orbit precessed through the diurnal cycle every 23 days. In this study, we used a regression formula developed by Chelle Gentemann in 2008, that uses the method described in Gentemann et al. (2003) but is derived using hourly ~ 5 km resolution, SST_{skin} data from the Meteosat Second Generation (MSG) Spinning Enhanced Visible and Infrared Imager (SEVIRI). The formula is as follows:

$$dSST(t, R, u) = f(t)e^{-0.27u} [a_1(R - R_0) + a_2(R - R_0)^2] \quad (1)$$

$$f(t) = [907.794 - 565.402\cos(\omega t) - 910.929\sin(\omega t) + 49.387\cos(2\omega t) + 292.148\sin(2\omega t) - 0.877\cos(3\omega t) - 68.033\sin(3\omega t) - 4.456\cos(4\omega t) + 2.214\sin(4\omega t) + 21.669\cos(5\omega t) - 2.117\sin(5\omega t)] \times 0.001 \quad (2)$$

where $f(t)$ represents the shape of the diurnal cycle through the day, t is local time (in hours since local solar midnight), u ($m s^{-1}$) is the wind speed at 10 m height, R ($W m^{-2}$) is solar shortwave insolation (SSI), $a_1 = 2.249 \times 10^{-3} K m^2 W^{-1}$, $a_2 = 2.014 \times 10^{-6} K m^4 W^{-2}$, $R_0 = 97 W m^{-2}$, and $\omega = 0.262 h^{-1}$. Since SSI is calculated from the solar zenith angle and solar constant, no allowance is made for cloud in this SEVIRI DV model. Also, as in the original Gentemann et al. (2003) formula, this model only calculates DV for solar heating $> 97 W m^{-2}$ conditions. Hereafter, we still use the term "CG03" as in several previous papers but it refers to this updated SEVIRI DV model.

Since the outputs of CG03 are dSST values only, to derive hourly $SST_{subskin}$ for comparison with MTSAT-1R $SST_{subskin}$ observations, it is necessary to add SST_{fnd} to hourly CG03 dSST values. The SST_{fnd} inputs are from the Regional Australian Multi-Sensor Sea surface temperature Analysis (RAMSSA; Beggs et al., 2011) developed at the Bureau. In this way, the hourly $SST_{subskin}$ is the sum of the input daily RAMSSA SST_{fnd} and hourly CG03 dSST values. Hourly 10 m wind speeds are obtained from the Australian Community Climate and Earth-System Simulator-Regional (ACCESS-R; Puri et al., 2013) 24 h forecasts. Details of the input data sets are in section 2.2.

2.1.2. The ZB05 Model

Zeng and Beljaars (2005) proposed a prognostic skin SST DV scheme. In ZB05, with prior knowledge of the wind speed, surface fluxes, and SST_{fnd} data, SST_{skin} can be calculated in two steps: the cool skin effect scheme and the warm layer profile scheme. For the cool skin effect, the equation is

$$T_s - T_{-\delta} = \frac{\delta}{\rho_w c_w k_w} (Q + R_s f_s) \quad (3)$$

where T_s is SST_{skin} , δ is the depth of the cool skin layer, $T_{-\delta}$ is $SST_{subskin}$, ρ_w is the sea water density with subscript w denoting sea water, c_w is the heat capacity of sea water, k_w is the molecular thermal conductivity, Q is the sum of the surface latent and sensible heat fluxes and the net longwave radiation defined as positive downward, R_s is surface net solar radiation, and f_s is the fraction of solar radiation absorbed in the sublayer.

The warm layer, the top of which is the bottom of the cool skin layer, is typically several meters deep (Donlon et al., 2007). The rate of change of temperature difference across this layer is

$$\frac{\partial(T_{-\delta} - T_{-d})}{\partial t} = \frac{Q + R_s - R_{(-d)}}{d \rho_w c_w v / (v+1)} - \frac{(v+1) k_{u,w}}{d \phi_t \left(\frac{d}{L}\right)} (T_{-\delta} - T_{-d}) \quad (4)$$

where d is the warm layer depth (WLD), T_{-d} is input SST_{fnd} , v is an empirical parameter which depends on d , $k = 0.4$ is the Von Karman constant, u_{*w} is the friction velocity in sea water via which wind speed enters, and L is the Monin-Obukhov length. On the right-hand side of equation (4), the first term is the rate of

heating due to surface fluxes and the second term is the cooling due to turbulent entrainment. The change in temperature at a particular time step is found from the discretized version of equation (4). ϕ_t is the stability function:

$$\phi_t\left(\frac{-z}{L}\right) = \begin{cases} 1 + 5\frac{-z}{L} & \text{for } \frac{-z}{L} \geq 0 \\ \left(1 - 16\frac{-z}{L}\right)^{-0.5} & \text{for } \frac{-z}{L} < 0 \end{cases} \quad (5)$$

In the TWP+ project, the input SST_{fnd} for ZB05 are also from RAMSSA. All meteorological variables are from ACCESS-R 24 h forecasts (see section 2.2.2). A constant WLD d of 3 m is assumed, as the DV of ocean temperature is often small at around this depth. Correspondingly, v is 0.3 (Zeng & Beljaars, 2005). The output of ZB05 is SST_{skin} (with cool skin scheme) or $SST_{subskin}$ (without cool skin scheme). DV is then calculated from SST_{skin} or $SST_{subskin}$.

2.1.3. The ZB + T Model

Takaya et al. (2010) proposed two refinements to the ZB05 warm layer scheme, including the modification of a Monin-Obukhov similarity function for stable conditions and the introduction of mixing enhancement by the Langmuir circulation. The modified stability function is

$$\phi_t\left(\frac{-z}{L}\right) = \begin{cases} 1 + \frac{5\frac{-z}{L} + 4\left(\frac{-z}{L}\right)^2}{1 + 3\frac{-z}{L} + 0.25\left(\frac{-z}{L}\right)^2} & \text{for } \frac{-z}{L} \geq 0 \\ \left(1 - 16\frac{-z}{L}\right)^{-0.5} & \text{for } \frac{-z}{L} < 0 \end{cases} \quad (6)$$

Also, under stable conditions the effect of the Langmuir circulation is included in the turbulent diffusivity, which transforms equation (4) to

$$\frac{\partial(T_{-\delta} - T_{-d})}{\partial t} = \frac{Q + R_s - R_{(-d)}}{d\rho_w c_w v / (v+1)} - \frac{(v+1)ku_{*w}}{d\phi_t\left(\frac{d}{L}\right)(u_{*w}/u_s)^{1/3}}(T_{-\delta} - T_{-d}) \quad (7)$$

where u_s is the surface Stokes velocity. We hereafter refer to the ZB05 scheme with these adjustments as the ZB + T scheme.

In this study, similarly as for CG03 and ZB05, the input SST_{fnd} for ZB + T are from RAMSSA and the meteorological inputs are from ACCESS-R. The output of ZB + T is also SST_{skin} (with cool skin scheme) or $SST_{subskin}$ (without cool skin scheme). DV is further calculated.

2.1.4. The GC2 Model

The GC2 model was released in March 2014. It is the latest configuration of the Met Office Unified Model, composed of component configurations Global Atmosphere 6.0 (GA6.0), Global Land 6.0 (GL6.0), Global Ocean 5.0 (GO5.0), and Global Sea Ice 6.0 (GSI6.0; Williams et al., 2015). In this study, we focus on the air-sea coupled components. The GA6.0 component has 85 levels up to 85 km height and ~ 60 km horizontal resolution, and the GO5.0 has 75 levels with 1 m resolution in the upper 10 and 2 m resolution between 10 and 20 m, and $0.25^\circ \times 0.25^\circ$ horizontal resolution at the equator (Megann et al., 2014). A 5 day hindcast is initialized at 0:00Z on each day in the 4 month period from the operational NWP atmospheric analysis and the FOAM (Fast Ocean Atmosphere Model; Blockley et al., 2014) ocean analysis. After initialization, the model is left to run freely for the forecast period (no observations are further assimilated). The first-day hindcast is used in this study, which is considered to be of better quality than the other four days since the model starts to lose skill as it evolves with time.

Although the vertical resolution in the GO5.0 global ocean model is high, the warm layer and cool skin will not be explicitly resolved. For this study, a DV scheme that consists of a warm layer and a cool skin scheme has been added to the coupled model. The warm layer scheme is the same as in ZB05, and the cool skin scheme is based on Artale et al. (2002). As a coupled model, GC2 uses the fourth ocean layer temperature, at a depth of 3.86 m, as input SST_{fnd} in the DV scheme. Unlike the other DV models in this paper, in GC2 the input SST_{fnd} is not from an analysis, but is generated dynamically by the ocean submodel. However, it should be noted that this ocean submodel is initialized by the FOAM ocean analysis, which assimilates similar satellite and in situ SST data to the RAMSSA SST analysis over the TWP region.

Table 1
Summary of the Input SST_{fnd} , Meteorological Variables, and DV Outputs of All Four Models

Model	SST_{fnd}	Meteorological Inputs	$SST_{skin}/SST_{subskin}$ DV?
CG03	RAMSSA	ACCESS-R outputs	N/Y
ZB	RAMSSA	ACCESS-R outputs	Y/Y
ZB + T	RAMSSA	ACCESS-R outputs	Y/Y
GC2	GC2 $SST_{3.86m}$	Within the model	Y/Y

Note. For the last column, Y (yes) and N (no) indicate whether or not the model can generate DV for the skin or subskin layer (see section 2.3).

Similar to ZB05 and ZB + T, the output of GC2 is also SST_{skin} or $SST_{subskin}$ and DV can be further calculated. The SST_{skin} is then coupled to the atmosphere model during the hourly coupling exchanges.

Table 1 summarises the models in terms of the input SST_{fnd} , meteorological variables, and the DV outputs.

2.2. Data Sets

2.2.1. MTSAT-1R Data

The reference observation SST data used in this study are the hourly, 0.05° , v3 MTSAT-1R data (Appendix A), produced by the Bureau over the TWP+ domain and period especially for the TWP+ project. The MTSAT-1R geostationary satellite was launched in 2005 by the Japan Aerospace Exploration Agency (JAXA) on behalf of the Japan Civil Aviation Bureau and the Japan Meteorological Agency. Only the highest quality data (quality level 5) are used.

In addition, all measurements that are within 15 km of the coast are discarded to remove possible terrestrial effects. This data set has been described and validated against both drifting buoy data and Advanced Along-Track Scanning Radiometer (AATSR) SST data in Zhang et al. (2016a). An overall zero bias and 0.73 K standard deviation was found from the in situ validation over the TWP+ domain and study period.

2.2.2. ACCESS-R Outputs

All meteorological data, used in this study as inputs for the CG03, ZB05 and ZB + T models, are average, hourly, 0.375° resolution outputs from the ACCESS-R model 24 h forecasts. This regional NWP model was developed by the Bureau in collaboration with the Commonwealth Scientific and Industrial Research Organisation (CSIRO) and the UK Met Office, and is based on the Met Office's Unified Model. It can predict a series of surface parameters, including wind speed, wind stress, shortwave and longwave radiative fluxes, and latent and sensible heat fluxes (Puri et al., 2013). The ACCESS-R wind speed and daily maximum SSI have been compared with the European Centre for Medium-Range Weather Forecasts (ECMWF) reanalysis SSI data and Cross-Calibrated Multiplatform (CCMP) sea surface wind data, respectively, over the Australian North-West Shelf from 2010 to 2014 (Wang & Zhang, 2017). Wang and Zhang (2017) found that these variables have good quality. ACCESS-R SSI and winds have a 0.85 and 0.88 correlation coefficient with ECMWF SSI and CCMP winds, respectively (Wang & Zhang, 2017).

2.2.3. RAMSSA Data

The Bureau's operational daily, $1/12^\circ$ resolution, RAMSSA SST analysis data (Beggs et al., 2011) have been used as direct SST_{fnd} inputs for the CG03, ZB05, and ZB + T models (Table 1). This data set is a level 4 gap-free SST analysis produced using optimal interpolation. To minimize the DV effects, the RAMSSA system ingests SST data from infrared and microwave sensors on polar-orbiting satellites, and from in situ measurements, only under conditions with ACCESS-R forecast 10 m wind speeds $> 2 \text{ m s}^{-1}$ during the night and $> 6 \text{ m s}^{-1}$ during the day. A much larger domain is covered (70°S – 20°N , 60°E – 170°W) than the TWP+ study region. A 0.42 K standard deviation is obtained when comparing RAMSSA SSTs with independent buoy SST observations for the period 1 October 2007 to 31 March 2008 (Beggs et al., 2011). More details about this data set can be found in Beggs et al. (2011) and Zhang et al. (2016b).

2.3. Method

In this study, the DV of $SST_{subskin}$, rather than SST_{skin} , are analysed for two reasons. First, although the sensor on-board MTSAT-1R measures temperature at $\sim 10 \mu\text{m}$ (classified as SST_{skin}), the radiances received were first regressed against drifting buoy SST observations at 20–30 cm depth to produce an approximation to $SST_{subskin}$. Then, a 0.17 K constant, rather than a physical cool skin effect scheme, was subtracted to produce SST_{skin} (Beggs et al., 2013). As a result, we obtain an approximation of DV for $SST_{subskin}$ when calculating the DV cycle from MTSAT-1R data. The second reason is that the CG03 model does not include the cool skin effect and can therefore only produce DV for $SST_{subskin}$. Like MTSAT-1R, the SEVIRI SST_{skin} data (Donlon et al., 2009) used to derive the CG03 model are also calculated as $SST_{subskin}$ minus a constant (in this case, 0.2 K). Therefore, DV is actually from $SST_{subskin}$. All the other models, namely ZB05, ZB + T and GC2, can produce both SST_{skin} (with a cool skin scheme) and $SST_{subskin}$ (without a cool skin scheme), and the latter is used in this study.

For satellite observations, foundation SSTs are normally estimated from an average of the predawn, or nighttime observed subskin SSTs (e.g., Deschamps & Frouin, 1984; Karagali et al., 2012; Zhang et al., 2016a).

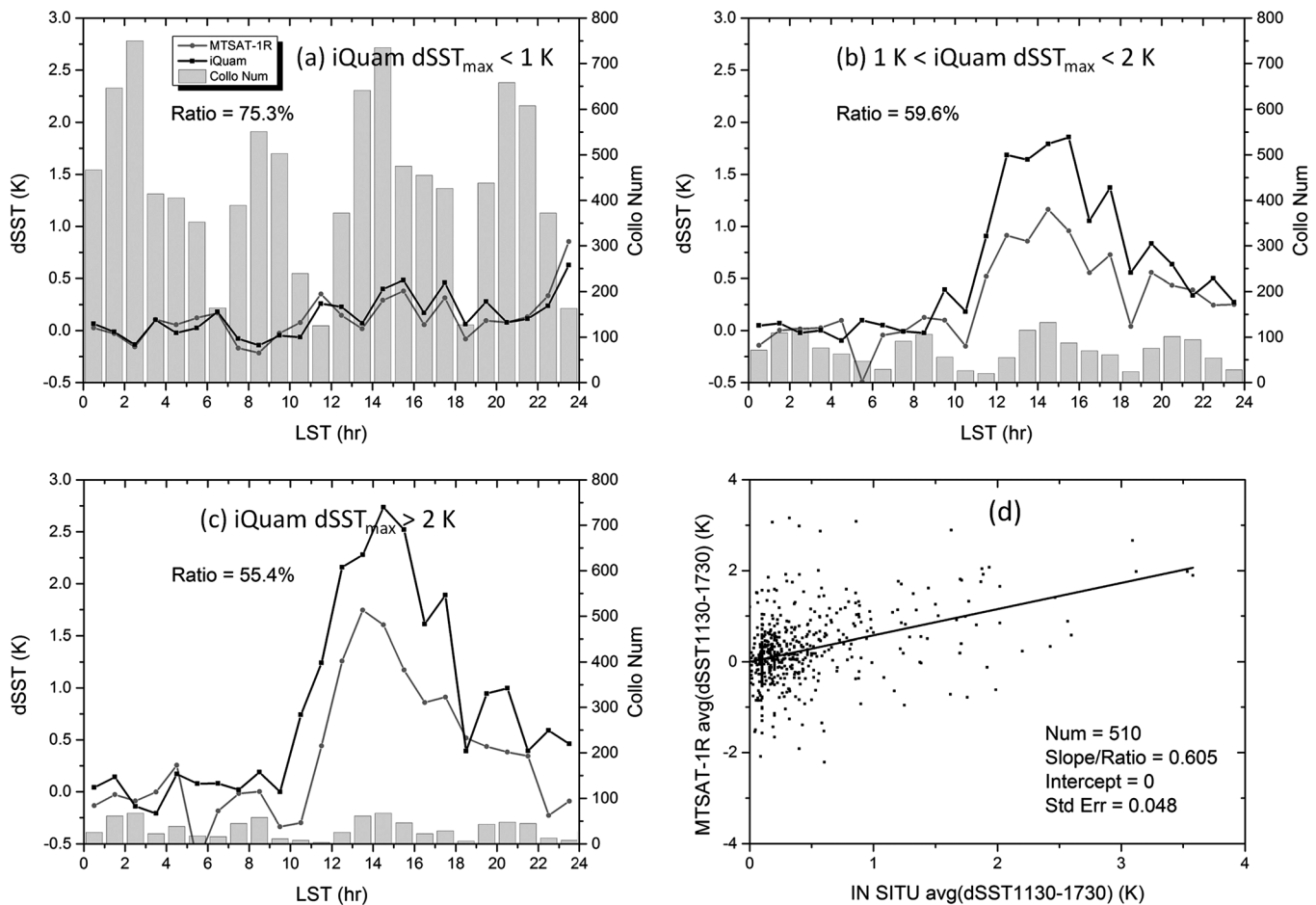


Figure 1. Sensitivity of MTSAT-1R to iQuam drifting buoy data for (a) iQuam dSST_{max} < 1 K; (b) 1 K < iQuam dSST_{max} < 2 K; and (c) iQuam dSST_{max} > 2 K. The ratio is calculated as the average MTSAT-1R dSSTs between 11:30 and 17:30 LST divided by the average drifting buoy dSSTs over the same time period. The grey columns are the collocation numbers over each hour. (d) A total least square fit for all the collocated average dSSTs.

Since the infrared sensor on-board MTSAT-1R cannot sense SST under cloud, to employ as much data as possible, we use the average of the early morning subskin SSTs, i.e., 00:30–05:30 LST, as an estimate of MTSAT-1R SST_{fndr}, called in this study the “constructed SST_{fndr}.” For the models, SST_{subskin} data are first calculated based on their direct SST_{fnd} inputs and each DV scheme. To keep consistent with MTSAT-1R, modeled constructed SST_{fnd} is also the average of the early morning subskin SSTs, i.e., 00:30–05:30 LST. Modeled SST_{subskin} DV cycles are subsequently recalculated from the constructed SST_{fndr}. Previous studies have shown that there is no significant difference in terms of DV amount or seasonal and spatial distribution through using the average early morning rather than predawn SSTs (e.g., Karagali et al., 2012; Karagali & Høyer, 2014).

Finally, data from all sources are averaged into bins with the lowest GC2 resolution $0.833^\circ \times 0.555^\circ$. Then all data are collocated both spatially and temporally under one condition: the number of MTSAT-1R measurements must be ≥ 15 out of 24 h within a local day at each grid cell to ensure that the MTSAT-1R DV cycle is properly represented. One drawback of this filter is that bias is introduced because more clear days than partly cloudy days are taken into consideration. However, collocating with other DV model and meteorological data reduces the sampling bias. In total, 169,004 collocations are obtained, on which all analyses are based.

3. Sensitivity of v3 MTSAT-1R SST Data

As the reference observational data, the v3 MTSAT-1R SST measurements are assumed to be of high quality and able to properly reflect DV cycles. Traditional in situ validation has illustrated the overall good performance of v3 MTSAT-1R data in Zhang et al. (2016a), as also mentioned in section 2.2.1. However, this cannot

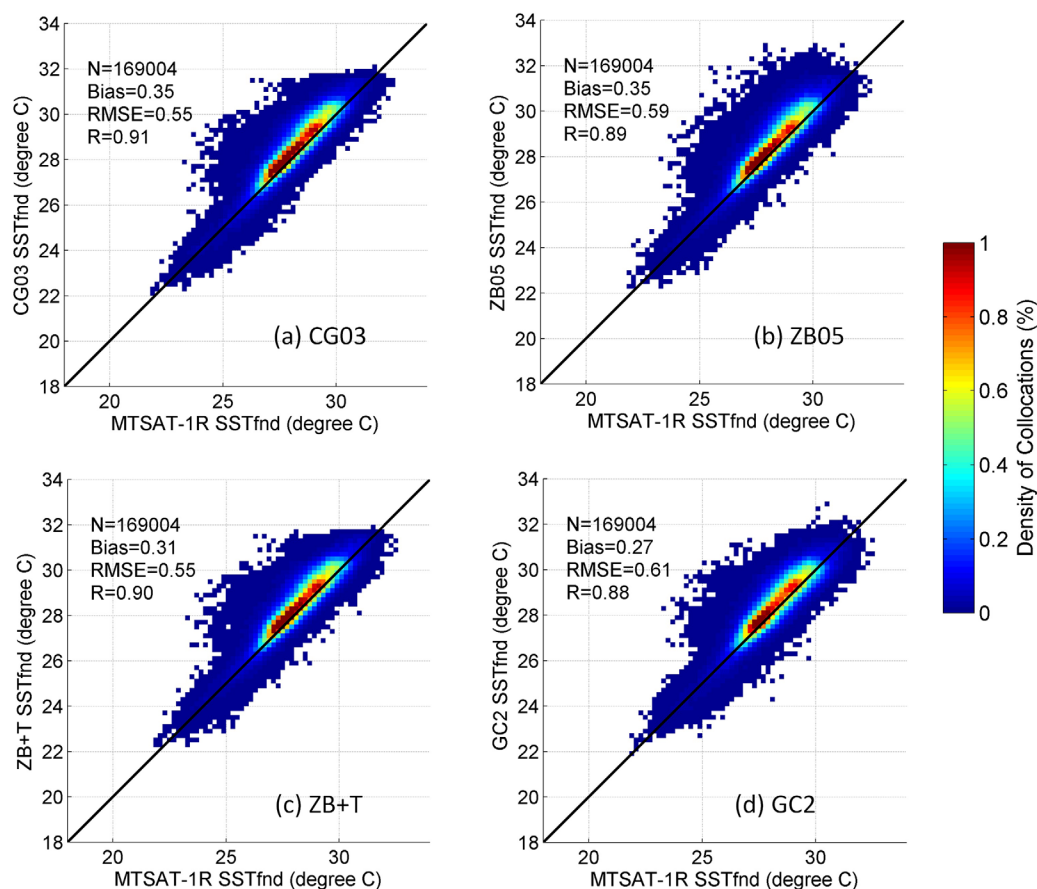


Figure 2. Density of the collocations between the modeled and MTSAT-1R constructed SST_{fnd} (averaged 00:30–05:30 LST SST_{subskin}). N is the number of collocations, RMSE is the root-mean-square error, and R is the correlation coefficient. The units for bias and RMSE are K.

guarantee its ability to resolve DV cycles. In addition to the typical validation criteria of the mean bias and standard deviation, Merchant et al. (2009, 2013) proposed that the sensitivity of SST retrievals to true changes in surface temperature should also be considered, especially in DV studies. They argue that different satellite retrievals of SST might have comparably similar, good qualities when validated against SST variability across large scales, yet have very different sensitivity to real changes in SST at scales that are small in space (such as across fronts) or time (such as subdaily variability). Sensitivity of 1.0 means that the full amplitude of the diurnal cycle in SST is captured by the satellite retrieval, which is clearly desirable for the purpose of this paper. The usefulness of sensitivity when assessing SST retrieval methods has been supported by the algorithm comparisons of Petrenko et al. (2014). Therefore, in this study, prior to validating the DV models against the v3 MTSAT-1R data set, we conduct a sensitivity test by comparing the diurnal cycles observed from MTSAT-1R and drifting buoys. Drifting buoys are considered suitable for this sensitivity analysis as they measure SST at around 20–30 cm depth, and for DV amplitudes > 1 K, the temperatures at these depths are relatively similar to those measured immediately below the cool skin (Gentemann et al., 2009). Moorings in the study region measure SST at ~ 1 m depth, and at high DV amplitudes the difference in temperature from the surface to 1 m is substantial, so mooring data were not used for the analysis. The drifting buoy SST data used in this section were obtained from the NOAA In Situ Quality Monitor version 2 (iQuam; Xu & Ignatov, 2014; <https://www.star.nesdis.noaa.gov/sod/sst/iquam/v2/data.html>). In iQuam v2 data, since the in situ SST values are quality controlled against foundation SST analyses that are regarded as “true” values, measurements with large DV signals are possibly flagged as low quality, since the foundation SSTs are free of DV. Therefore, in this study, we did not apply any of the quality_level flags to the iQuam data. The methods to construct SST_{fnd} estimates from both MTSAT-1R and iQuam data are the same, both being the average of early morning (00:30–05:30 LST) measurements (see section 2.3). When establishing the collocations between in situ and satellite data,

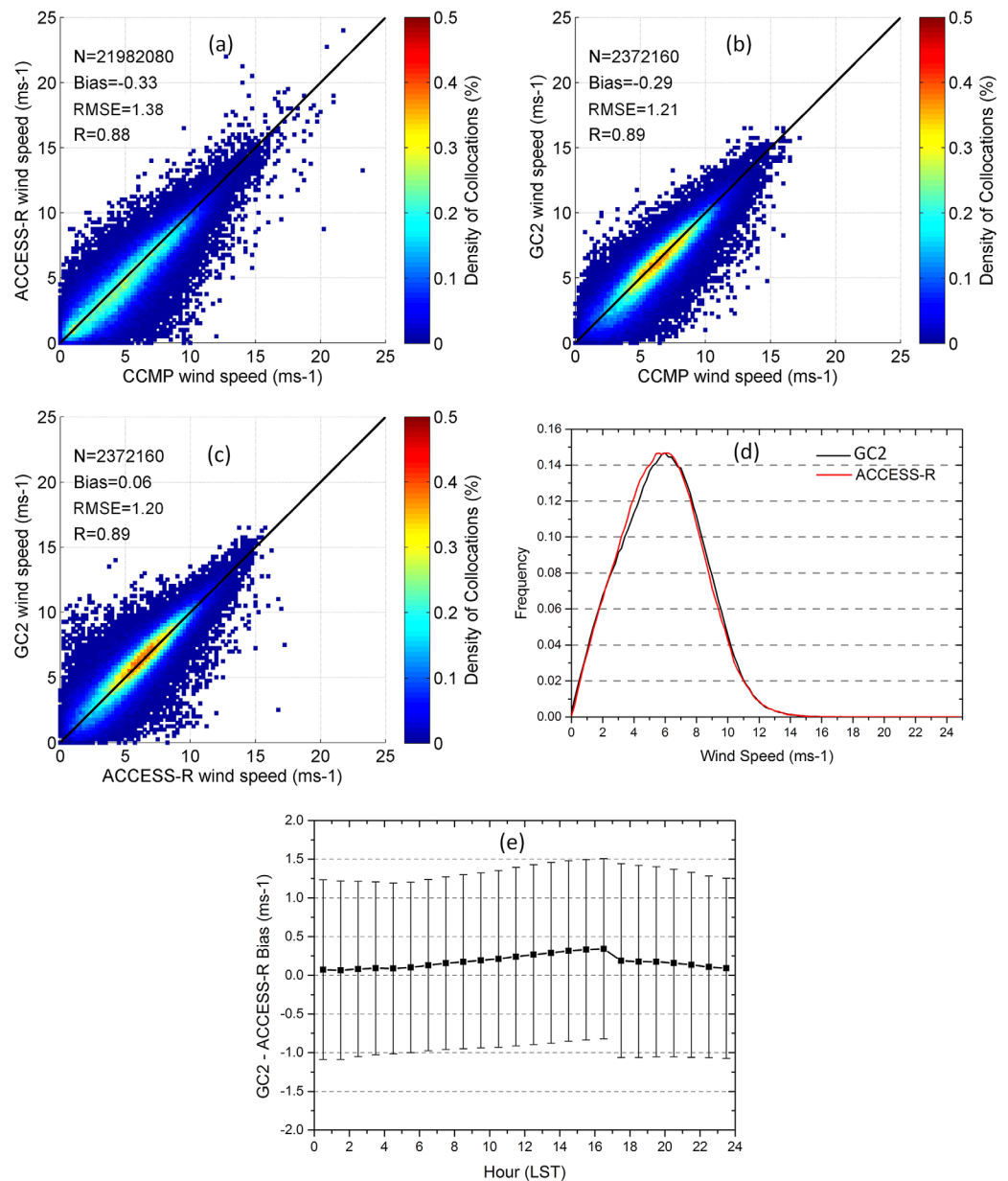


Figure 3. Validation of wind speed data from (a) ACCESS-R, and (b) GC2 against CCMP wind data, (c) intercomparison between GC2 and ACCESS-R winds, (d) PDF (probability density function) plot of GC2 and ACCESS-R winds, and (e) biases \pm STD between GC2 and ACCESS-R winds as a function of local time.

the following filters are applied: (1) $0.15^\circ \times 0.15^\circ$ spatial window and ± 30 min temporal window; (2) any collocation with absolute difference > 10 K is discarded; (3) for both DV cycles, there should be at least three measurements between 00:30 and 05:30 LST, and at least four between 11:30 and 17:30 LST. As suggested by Merchant et al. (2013), we calculate the sensitivity ratio as the average MTSAT-1R dSSTs between 11:30 and 17:30 LST divided by the average drifting buoy dSSTs over the same time frame.

As shown in Figure 1, MTSAT-1R noticeably underestimates the iQuam DV amplitudes under different DV conditions. When iQuam dSST_{max} is < 1 K, the ratio is the highest, reaching 75.3%, although this ratio has high uncertainty because it is a ratio of small values. The ratios for iQuam dSST_{max} > 1 K are between 55% and 60%, although the collocations are much fewer. In Figure 1d, a total least square linear fit, accounting errors from both axes, is added to all the collocations. Forcing the fitting line to cross the origin results in a slope (\pm statistical fitting uncertainty) of 0.60 (± 0.05). This is used as an overall estimate of the MTSAT-1R

Table 2

Statistical Analyses of the Collocations Between the Modeled $dSST_{max}$ and the Original MTSAT-1R $dSST_{max}$ As Well As the Scaled MTSAT-1R/Ratio $dSST_{max}$ Data

	Bias (K)		RMSE (K)		R	
	MTSAT-1R $dSST_{max}$	MTSAT-1R/ ratio $dSST_{max}$	MTSAT-1R $dSST_{max}$	MTSAT-1R/ ratio $dSST_{max}$	MTSAT-1R $dSST_{max}$	MTSAT-1R/ ratio $dSST_{max}$
CG03	−0.16	−0.68	0.54	0.85	0.48	0.48
ZB05	0.14	−0.38	0.81	0.97	0.45	0.45
ZB + T	−0.20	−0.71	0.62	0.87	0.46	0.46
GC2	0.72	0.20	0.98	1.08	0.46	0.46

Note. Statistics include average bias of the modeled $dSST_{max}$ minus MTSAT-1R $dSST_{max}$ (and MTSAT-1R/ratio $dSST_{max}$), root-mean-square errors (RMSE), and correlation coefficient (R).

sensitivity. This result is similar to the findings for this geographical area in Merchant et al. (2009; their Figure 1c), where the sensitivity estimate was based on radiative transfer simulation rather than empirical data. Although Merchant et al. (2009) were investigating a different satellite sensor, the Advanced Very High Resolution Radiometer (AVHRR), the thermal channels and principle of retrieval are similar (see Appendix A for the v3 MTSAT-1R retrieval algorithm).

Given an estimated mean sensitivity of 0.60, an observed diurnal cycle of 1 K in MTSAT-1R data would imply a true diurnal amplitude around 1.7 K. This difference is significant when using MTSAT-1R data for validating DV models. Therefore, for the following analyses, we will compare the four DV models against both the original MTSAT-1R and the sensitivity-scaled estimated “MTSAT-1R/ratio” (overall ratio 0.60) DV, given that the drifting buoy measurements with large diurnal amplitudes are too few for conclusive comparison. The sensitivity-scaled “MTSAT-1R/ratio” DV is considered most realistic in this study.

4. Results

4.1. Preanalyses

Due to the different SST_{fnd} sources and meteorological inputs between GC2 and the other DV models, we believe that before we directly analyse the DV amplitudes, it is necessary to assess the extent that these different inputs impact the modeled DV results. The constructed SST_{fnd} and wind speed data are compared here.

When comparing the constructed modeled SST_{fnd} and MTSAT-1R SST_{fnd} values (Figure 2), it was found that out of 169,004 collocations, most SST_{fnd} data fall between 27 and 30°C. As expected, the statistics for CG03, ZB05, and ZB + T SST_{fnd} are similar since the models have the same RAMSSA SST_{fnd} data inputs (Figures 2a–2c). Compared to MTSAT-1R, SST_{fnd} data from these three models all slightly overestimate the observations with a positive mean bias of 0.31–0.35 K. It should also be noted that the 00:30–05:30 LST MTSAT-1R $SST_{subskin}$ data display a small warm bias (~ 0.1 K) when compared with collocated drifting buoy SST observations at 0.2 m depth over the same region and period (Zhang et al., 2016a). The positive mean bias in the modeled SST_{fnd} is partly due to RAMSSA SST_{fnd} being warmer than the average 00:30–05:30 LST MTSAT-1R $SST_{subskin}$ since in addition to nighttime observations, the RAMSSA system blends daytime and early evening SST observations under high wind speed conditions (Beggs et al., 2011). The root-mean-square errors (RMSEs) in Figures 2a–2c are all < 0.6 K. Encouragingly, the validation of GC2 SST_{fnd} data is very similar to the other three models with a marginally smaller positive bias (0.27 K) and higher RMSE (0.61 K; Figure 2d). There are three possible factors accounting for the warm bias of GC2 SST_{fnd} data: (1) the temperature at WLD used in the warm layer scheme is taken from the ocean model, and it might suffer from a warm bias. This, in turn, could produce a warm bias in the SST_{fnd} ; (2) there may be a positive surface shortwave radiation bias in most of the TWP+ region during the study period; (3) the FOAM SST analyses used to initialize the GC2 model might suffer from a positive bias. Determining the source of the GC2 SST_{fnd} warm bias source is beyond the scope of the present paper.

We then checked if the meteorological variables within the GC2 coupled model are similar to the ACCESS-R outputs by conducting a brief validation of the wind speeds from ACCESS-R and GC2 against CCMP ocean surface wind data (Atlas et al., 2011). According to Zhang et al. (2016a, 2016b), wind is the most important

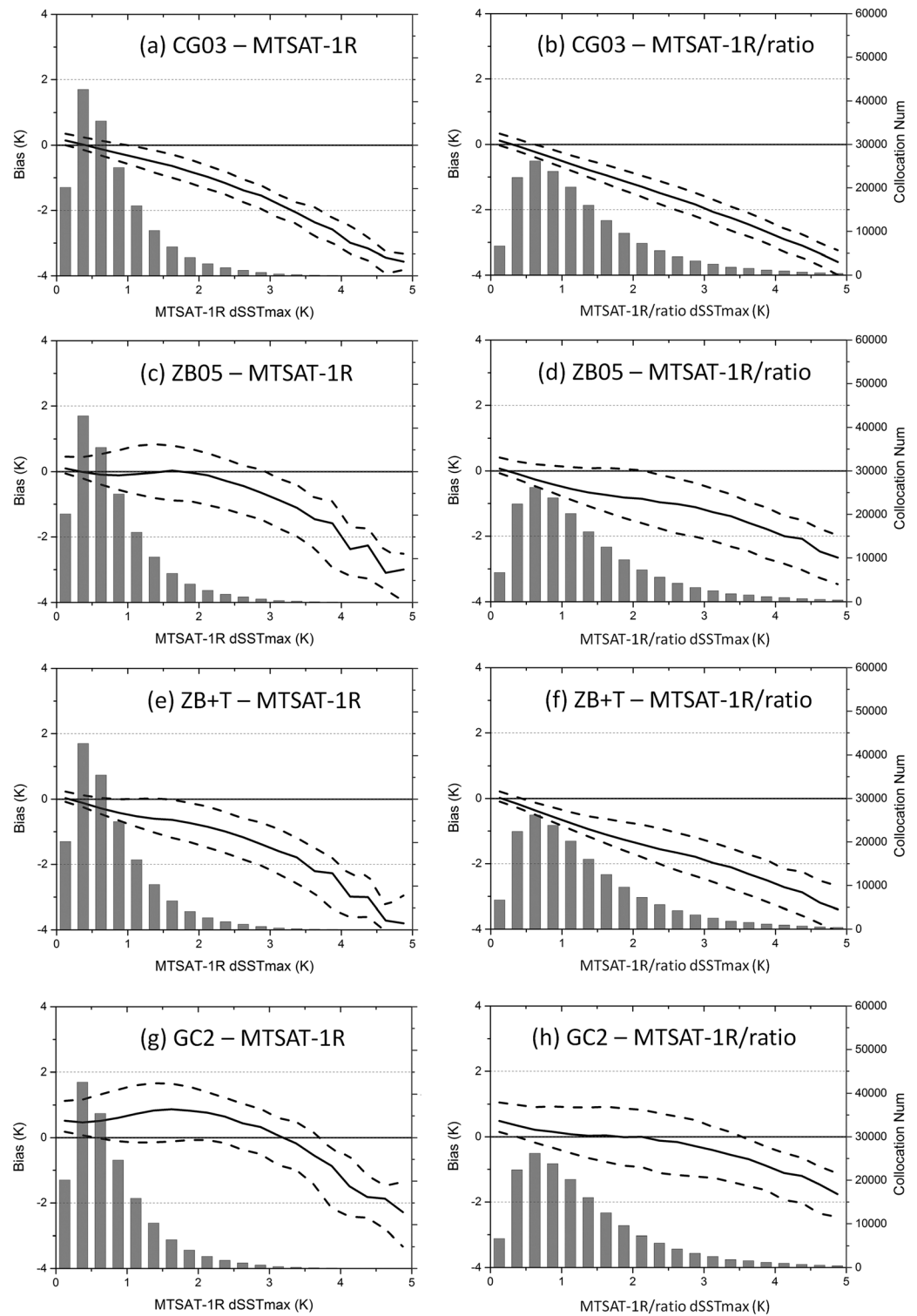


Figure 4. (left column) Biases in dSST_{max} of the collocations within different original MTSAT-1R dSST_{max} 0.25 K intervals through the whole 0–5 K range. The solid line is the median bias, and the bottom and top dashed lines are the 25th and 75th percentiles, respectively. The grey columns are the collocation numbers falling within each 0.25 K interval. (right column) As for the left column but for MTSAT-1R/ratio dSST_{max}.

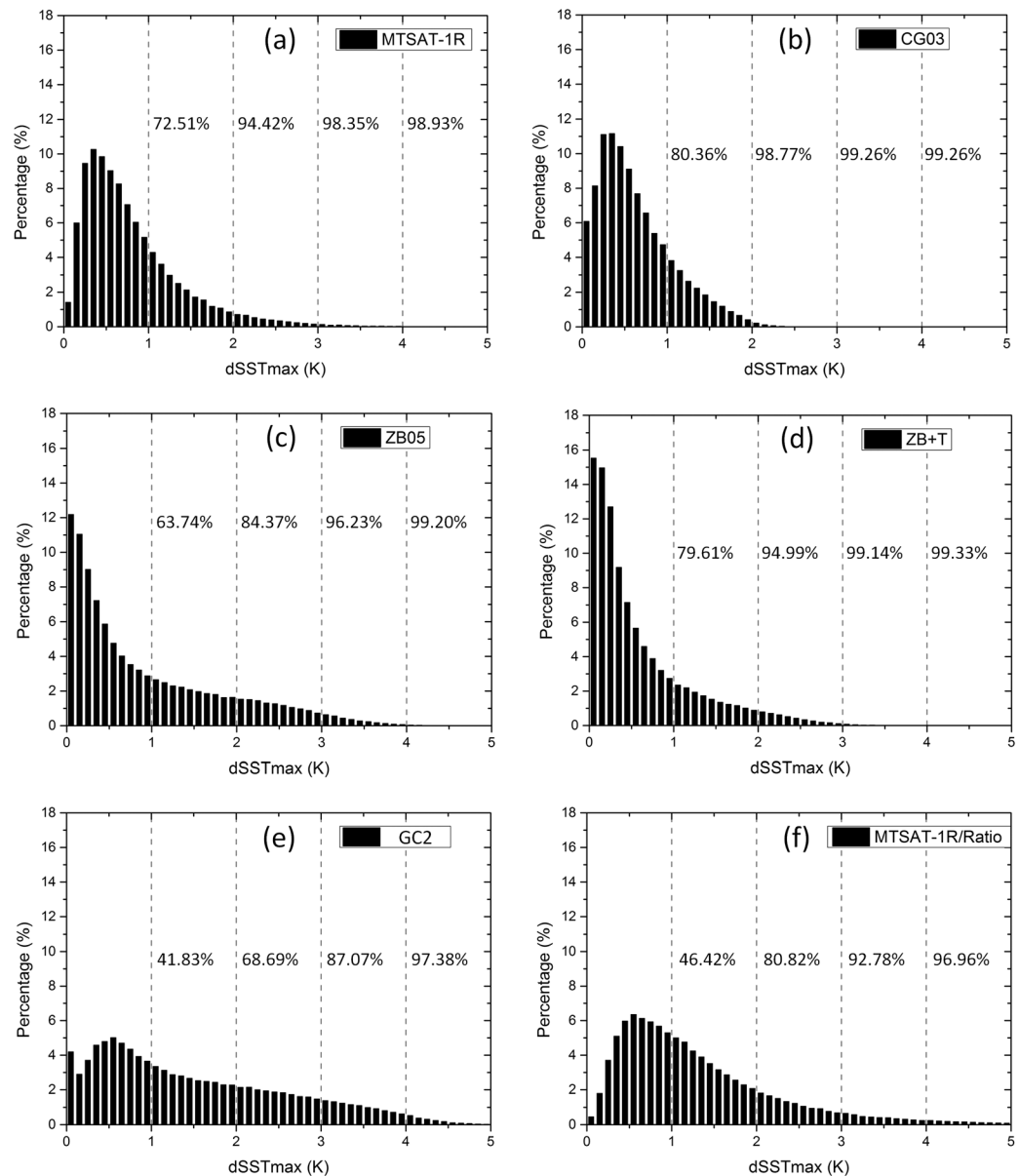


Figure 5. Distribution of dSST_{max} values from (a) original MTSAT-1R, (b) CG03, (c) ZB05, (d) ZB + T, (e) GC2, and (f) MTSAT-1R/ratio on 0.1 K intervals. The percentages are the accumulated proportions of dSST_{max} < 1 K, < 2 K, < 3 K, and < 4 K (the dashed reference lines).

meteorological factor regulating DV over the TWP region. Figures 3a and 3b indicate that the two wind data sets perform similarly well against CCMP winds, with $\sim 0.3 \text{ m s}^{-1}$ biases and ~ 0.9 correlation coefficients, and very similar probability density functions (Figure 3d). The average bias between the two wind data sets is very small, 0.06 m s^{-1} (Figure 3c). In Figure 3e, the biases between GC2 and ACCESS-R winds are plotted as a function of local times. Figure 3e shows no clear diurnal pattern in the biases.

Overall, the consistency in SST_{fnd} and meteorological inputs between the models gave us confidence to conduct further DV analyses.

4.2. Statistical Analyses

The maximum SST DV amplitudes during each solar day, dSST_{max}, from the models and MTSAT-1R were statistically compared (Table 2). This indicates that the agreement between all models and observations is relatively moderate with correlation coefficients between 0.45 and 0.48. This low correlation has been noted in

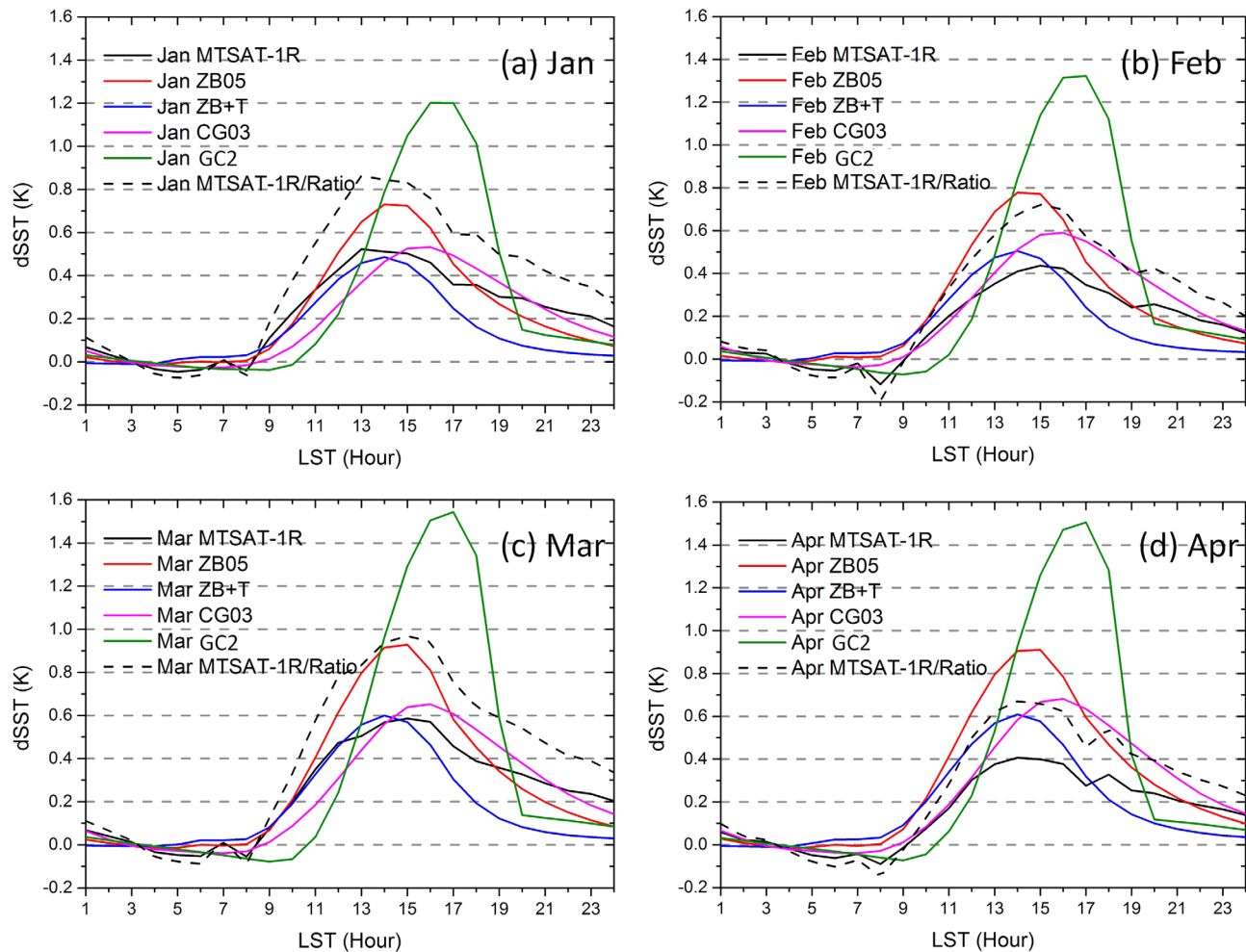


Figure 6. Monthly average DV cycles from MTSAT-1R and MTSAT-1R/ratio and four models for January–April 2010.

previous studies. For example, in Noh et al. (2011), a correlation of 0.31 between the modeled and buoy dSSTs is found, although it should be noted that their DV model, meteorological data and study domain differ from this TWP study. A possible reason might be that the surface forcing from the models does not match exactly with the real ocean, in spite of the resemblance in the large-scale weather pattern (Noh et al., 2011). Compared with the original MTSAT-1R data, CG03 and ZB + T $dSST_{max}$ data are slightly cold biased by -0.16 and -0.20 K, respectively. The $dSST_{max}$ from ZB05 is overestimated by 0.14 K, while the GC2 is significantly warm biased. However, when MTSAT-1R $dSST_{max}$ is scaled by the sensitivity ratio (0.60), the biases between the modeled $dSST_{max}$ and MTSAT-1R/ratio $dSST_{max}$ decline sharply. The negative bias in ZB05 (-0.38 K) is smaller than that for CG03 and ZB + T (-0.68 and -0.71 K). The GC2 has the smallest positive bias, being 0.2 K. All RMSEs have increased in the MTSAT-1R/ratio $dSST_{max}$ scenario. The RMSE is smaller in CG03 and ZB + T than in ZB05 and GC2.

The 25th/50th(median)/75th percentile biases between the modeled $dSST_{max}$ and MTSAT-1R $dSST_{max}$ at each 0.25 K interval for the 0–5 K range were calculated, together with the collocation numbers (Figure 4). Figure 4 shows that as DV becomes large, all the models have an increasingly negative $dSST_{max}$ bias. Compared with both the original MTSAT-1R and the sensitivity-scaled MTSAT-1R/ratio $dSST_{max}$, CG03 has the narrowest 25th to 75th percentile bias range, which is not surprising because this model has fewer inputs than ZB05, ZB + T or GC2. Apart from wind and SSI, inputs in these other DV models such as latent and sensible heat fluxes can possibly increase the output uncertainties. The cold bias in CG03 is nearly linear as MTSAT-1R or MTSAT-1R/ratio $dSST_{max}$ increases (Figures 4a and 4b). ZB05 has near zero median bias for the 0–2 K range but fails to properly estimate larger DV events (Figure 4c). ZB + T also

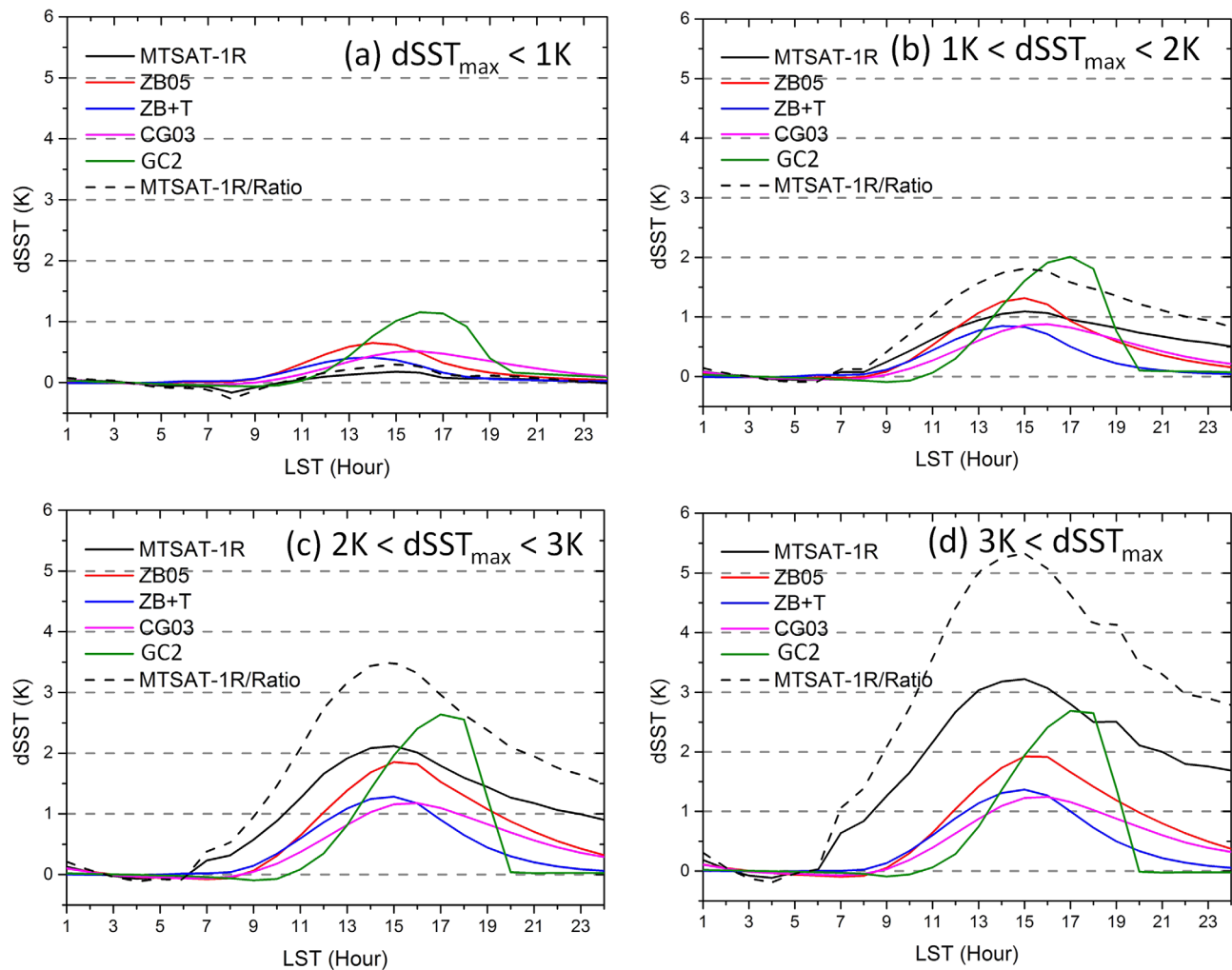


Figure 7. Four month average DV cycles from MTSAT-1R and MTSAT-1R/ratio and four models for different original MTSAT-1R $dSST_{max}$ conditions: (a) $dSST_{max} < 1$ K, (b) $1\text{ K} < dSST_{max} < 2$ K, (c) $2\text{ K} < dSST_{max} < 3$ K, and (d) $dSST_{max} > 3$ K.

underestimates $dSST_{max}$ for almost the whole range with a slightly narrower 25th/75th bias range than ZB05 (Figure 4e). GC2 displays a quite similar bias pattern to ZB05, largely due to the same warm layer scheme. However, instead of a near zero bias, GC2 has a positive median bias for $dSST_{max} < 3$ K (Figure 4g). Relative to the sensitivity-scaled MTSAT-1R/ratio $dSST_{max}$ data, biases in all models have become more negative to different extents at different DV ranges (Figures 4b, 4d, 4f, and 4h). GC2 now estimates $dSST_{max}$ values closest to MTSAT-1R/ratio $dSST_{max}$, as indicated by the near zero median bias for the 0–3 K range (Figure 4h).

Further analysis of the distribution of the MTSAT-1R and modeled $dSST_{max}$ values is shown in Figure 5. In the original MTSAT-1R data, for all the 0.1 K intervals, the largest percentage (10.3%) of $dSST_{max}$ values fall between 0.3 and 0.4 K (Figure 5a). There are 72.5% < 1 K and only 1.7% > 3 K. Shape-wise, the CG03 model best captures the original MTSAT-1R pattern, especially for $dSST_{max} < 1$ K, with the maximum percentage also observed between 0.3 and 0.4 K. When compared with the estimated MTSAT-1R/ratio $dSST_{max}$ (Figure 5f), the distribution closest to the MTSAT-1R/ratio now is the GC2 in terms of the overall shape, although GC2 has fewer 0–1 K and more 2–4 K $dSST_{max}$ values than the MTSAT-1R/ratio. This explains why the bias is still a positive 0.2 K (Table 2). ZB05 estimates a similar percentage for the 2–3 K range (11.86% in comparison with 11.96% in MTSAT-1R/ratio) but many more DV events < 1 K (63.74% compared with 46.42%). CG03 and ZB + T are severely underestimating the DV amplitudes for the whole 0–5 K range relative to MTSAT-1R/ratio.

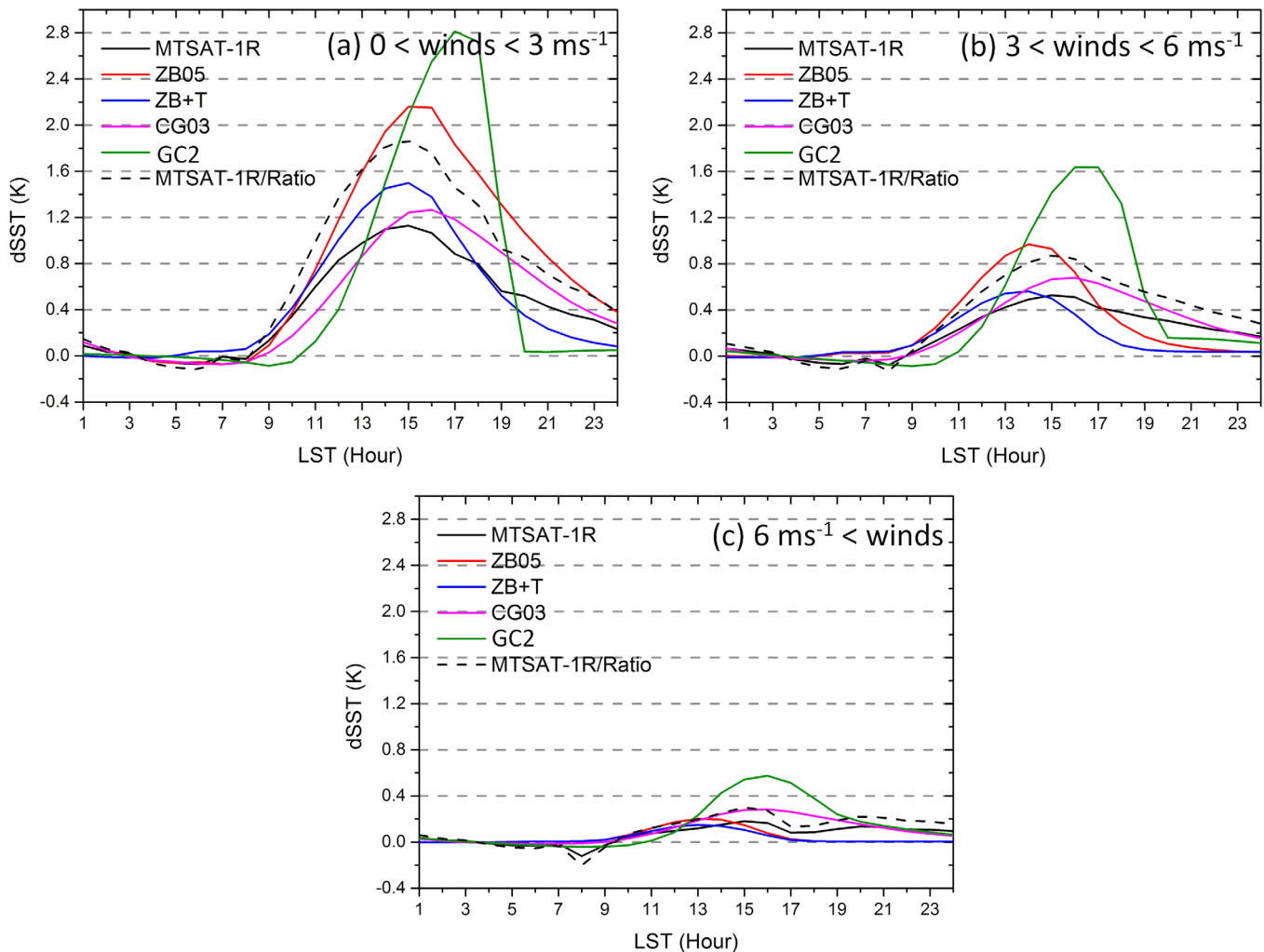


Figure 8. Four month average DV cycles from MTSAT-1R and MTSAT-1R/ratio and all four models for different wind speed conditions: (a) wind speed $< 3 \text{ m s}^{-1}$, (b) $3 \text{ m s}^{-1} < \text{wind speed} < 6 \text{ m s}^{-1}$, and (c) wind speed $> 6 \text{ m s}^{-1}$. Note that the 24 h forecast ACCESS-R 10 m wind speeds are given as an average at the end of each hour.

4.3. DV Cycles

Monthly average DV cycles calculated from the 4 months' collocations are plotted in Figure 6. The original MTSAT-1R data show that the average DV cycle has an amplitude of 0.4–0.6 K, typically peaking at 13–15 LST, but this value is increased to 0.8–0.9 K in the sensitivity-scaled MTSAT-1R data (Figure 6a). The average minimum temperatures are observed at 5–6 LST. Overall, CG03 and ZB + T agree well with the original MTSAT-1R pattern for most of the months. However, CG03 has its average DV cycle peak at 16 LST and ZB + T at 14 LST. The positive bias (relative to the original MTSAT-1R) in the average DV cycle amplitude in ZB05 is seen in every month, reaching 0.5 K in April. When compared with the sensitivity-scaled estimated MTSAT-1R/ratio DV cycles (dashed lines in Figure 6), CG03 and ZB + T underestimate DV amplitudes in all months except for April. ZB05 closely follows the estimated DV cycles in most months. However, in both original MTSAT-1R and sensitivity-scaled MTSAT-1R/ratio scenarios, the $dSST_{\max}$ values in GC2 are highly overestimated in all months, exceeding 1 K in April. In addition, the times when warming starts (~ 10 LST) and peaks (16–17 LST) in GC2 are also 1–3 h later than in the observations and other models. As a consequence, GC2 SSTs cool more rapidly up until 20 LST, after which the temperature stays relatively stable.

Figure 6 displays discontinuities in MTSAT-1R data at 6–8 LST and 17–19 LST. This feature has been observed and explained in Zhang et al. (2016a). To summarize, at local sunrise (6–8 LST in the study domain and period) and sunset (17–19 LST), there are fewer good quality MTSAT-1R SST data than at other times in

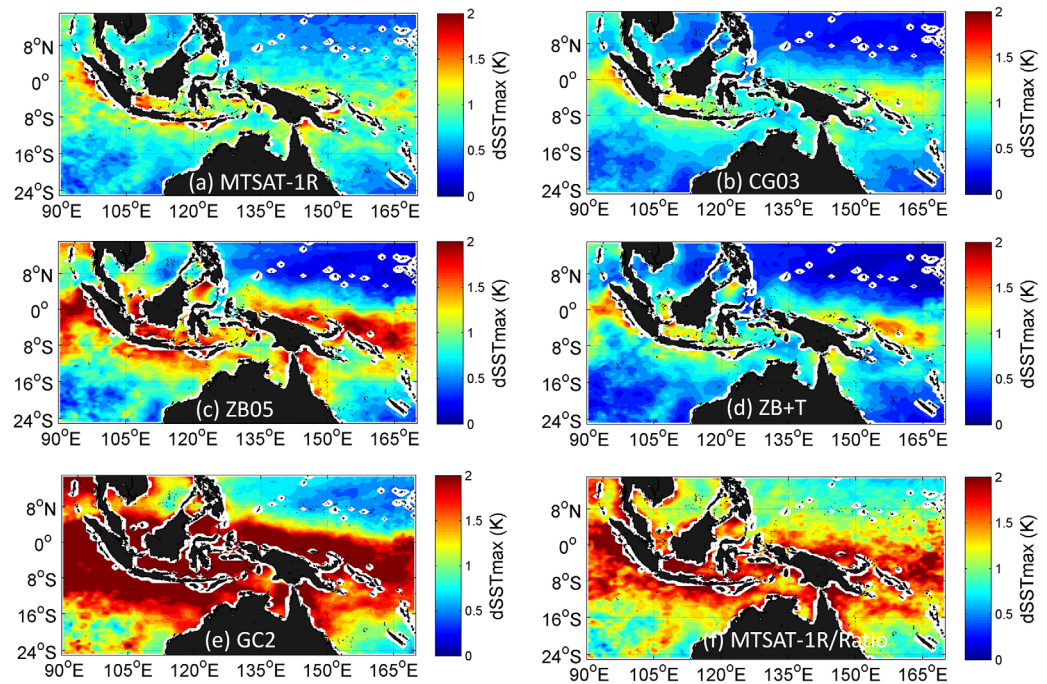


Figure 9. Spatial distribution of temporally averaged $dSST_{max}$ values over the study period from the (a) original MTSAT-1R, (b) CG03, (c) ZB05, (d) ZB + T, (e) GC2 model, and (f) MTSAT-1R/ratio.

the day due to the cloud screening method used in data production. The cloud mask depends on both visible and infrared channels in the daytime but only on infrared channels at night. Dusk and dawn are times when the visible channels cannot be used because of weak light, yet the $3.7 \mu m$ infrared channel is still contaminated by visible light. Consequently, there is greater uncertainty in the cloud detection, and more data during these times are therefore flagged as poor quality.

For each model, 4 month average DV cycles were plotted for different observed ranges of the original MTSAT-1R $dSST_{max}$: <1 K, $1-2$ K, $2-3$ K, and >3 K (Figure 7). For the original MTSAT-1R $dSST_{max} < 1$ K condition, only a 0.2 K average $dSST_{max}$ is observed in MTSAT-1R. The $dSST_{max}$ values for most of the models are between 0.4 and 0.6 K, except for GC2 at 1.1 K (Figure 7a). Little difference is observed in the MTSAT-1R/ratio scenario in Figure 7a. As MTSAT-1R $dSST_{max}$ increases to $1-2$ K, agreement between the original MTSAT-1R and CG03, ZB + T, and ZB05 is closest, in terms of both $dSST_{max}$ and the shape of the DV cycles (Figure 7b). However, when compared with scaled MTSAT-1R/ratio DV, all models, except for GC2, underestimate DV amplitudes by $0.5-1$ K. When the original MTSAT-1R $dSST_{max}$ is $2-3$ K (Figure 7c), CG03, ZB05, and ZB + T tend to underestimate the amplitudes, but GC2 has a positive bias of ~ 0.5 K. For MTSAT-1R $dSST_{max} > 3$ K (Figure 7d), all models underestimate the DV, especially for the scaled MTSAT-1R/ratio scenario. However, as shown in Figure 5, DV events > 3 K are quite rare.

Another observation should be noted. As MTSAT-1R $dSST_{max}$ increases, average $dSST$ at the end of the local day (23:30 LST) also increases in both the original and scaled MTSAT-1R data, reaching ~ 1.7 K and ~ 3 K (respectively) at 24 LST for $dSST_{max} > 3$ K (Figure 7d). This residual warmth over the TWP region has also been observed and explained in Zhang et al. (2016a). Fundamentally, two factors contribute to this difference. First, our SST_{ind} construction method (average of SSTs from 00:30 to 05:30 LST) determines that the $dSSTs$ approach 0 K in the early morning. The second is that for DV events $> 1-2$ K, the cooling before midnight cannot totally cancel the heating during the day unless there are sudden strong wind bursts causing significant turbulent mixing. This residual warmth has also been seen in other studies over different research domains, such as over the Atlantic Ocean in Karagali and Høyer (2014). It would appear that none of the models fully captures this residual warmth.

Further analysis of the modeled DV cycles was conducted under different daily average ACCESS-R wind speed conditions (Figure 8). We classified winds into three categories: low winds ($< 3 m s^{-1}$), moderate

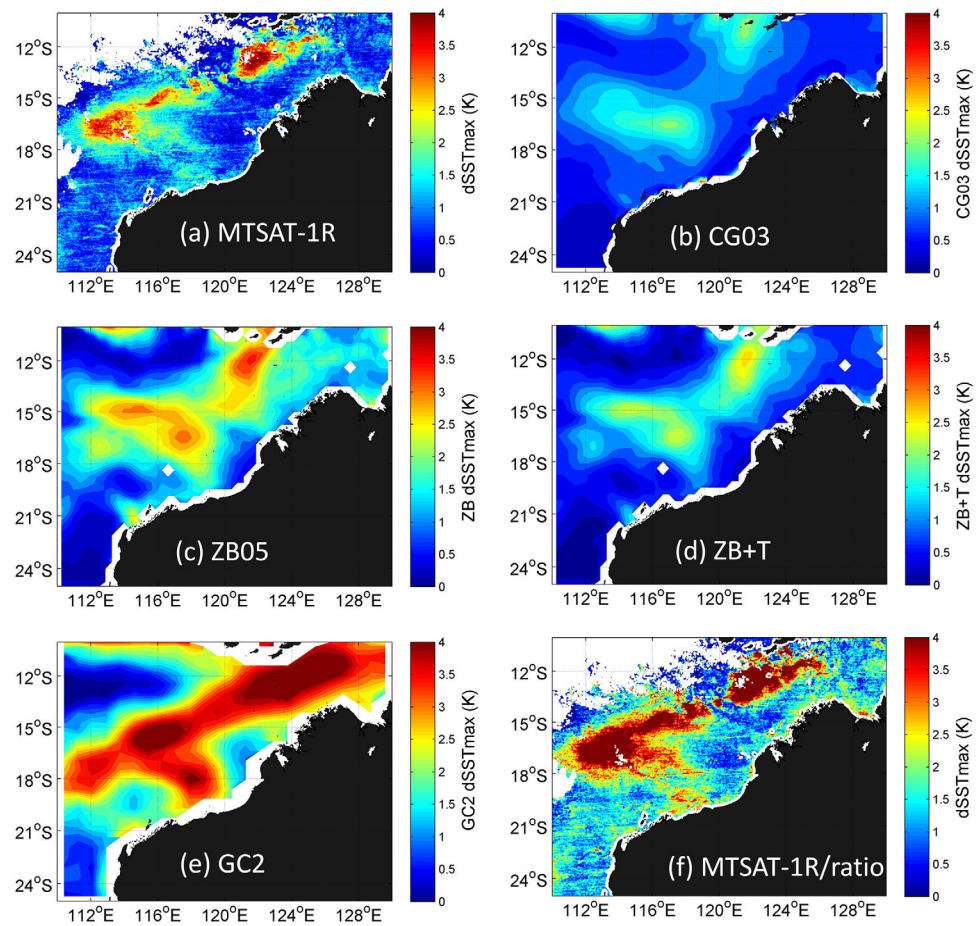


Figure 10. A DV event over the north-west Australian coast on 6 March 2010. Plots are the spatial distributions of $dSST_{max}$ from (a) original MTSAT-1R, (b) CG03, (c) ZB05, (d) ZB + T, (e) GC2, and (f) scaled MTSAT-1R/ratio data.

winds (between 3 and 6 m s^{-1}), and high winds ($> 6 \text{ m s}^{-1}$). Under low wind speed conditions, all models tend to overestimate the original MTSAT-1R $dSST_{max}$ ($\sim 1.1 \text{ K}$), especially GC2 by nearly 1.7 K (Figure 8a). When compared with the sensitivity-scaled MTSAT-1R/ratio data, ZB05 reflects the DV cycle relatively well, whereas CG03 and ZB + T underestimate the scaled MTSAT-1R data. For moderate winds, significantly smaller $dSST_{max}$ values are found for all models (Figure 8b), indicating the strong dependence of the models on wind speed. ZB05 appears to agree most closely with the sensitivity-scaled MTSAT-1R/ratio data. For high winds, no significant DV is observed except for GC2, which consistently overestimates $dSST_{max}$ values under all wind conditions (Figure 8c).

4.4. Spatial Distribution

Spatial distributions of temporally averaged daily $dSST_{max}$ values for each grid point, over the whole study period, for MTSAT-1R and all models were investigated (Figure 9). In Figure 9a, average DV signals $> 1 \text{ K}$ are observed along the coasts of Indonesia, New Guinea Island, and the surrounding islands within the 15°S – 8°N zone. When scaled by the sensitivity ratio, the DV amplitudes reach 1.5 – 2 K (Figure 9f). This DV distribution feature over the TWP domain has been observed in several previous studies (e.g., Stuart-Menteth et al., 2003; Weihs & Bourassa, 2014). CG03 and ZB + T are able to reflect the original MTSAT-1R distribution quite well, not only spatially but also amplitude wise. Compared with the scaled MTSAT-1R data (Figure 9f), ZB05 estimates the closest $dSST_{max}$ in terms of both spatial distribution and DV amplitudes. However, strongly overestimated $dSST_{max}$ values over a much larger region are observed in the GC2 model for both original MTSAT-1R and sensitivity-scaled MTSAT-1R/ratio scenarios.

Following investigations into the distribution of averaged $dSST_{max}$, we also conducted a case study of a strong DV event that occurred on 6 March 2010 over the north-west Australian coast (Figure 10). Original

MTSAT-1R indicates that $dSST_{max}$ can reach 4 K over a linear region to the north of the Australian north-west coast (Figure 10a). CG03 and ZB + T capture the $dSST_{max}$ spatial distribution quite effectively, but their amplitudes are much smaller (1.5–2 K). Better agreement is found in ZB05 with $dSST_{max}$ reaching 3.5 K. GC2 shows a much stronger amplitude over a larger region. However, when compared with the sensitivity-scaled MTSAT-1R/ratio (Figure 10f), GC2 is the only model capturing a similar DV range, while ZB05 underestimates by 0.5–1 K, and CG03 and ZB + T underestimates by 2–2.5 K.

5. Discussion and Conclusions

SST diurnal variability schemes are missing from many air-sea coupled, numerical weather prediction, and climate models. Including the effects of SST diurnal variation is expected to improve these models' performance (Marullo et al., 2016). SST data assimilation also requires incorporation of DV to meet the demand for increasingly high accuracy. Under such circumstances, selecting an appropriate DV model from quite a number of different models is a key preliminary step. Although the DV models have all been validated against in situ or satellite observations when proposed or in later papers, to the best of the authors' knowledge, there has been no research conducted to intercompare several DV models over the Tropical Warm Pool. In hope of providing useful information for air-sea coupled, NWP or climate modelers, this study has evaluated and intercompared four DV models, including the empirical CG03 model, physical ZB05 and ZB + T models, and air-sea coupled GC2 model, against the Bureau of Meteorology's v3 MTSAT-1R data as a contribution to the GHRSSST TWP+ project.

Before the v3 MTSAT-1R data were used to evaluate the models, their sensitivity to changes in "true" SSTs (iQuam drifting buoy SSTs in this study) was tested. According to Merchant et al. (2009, 2013), the sensitivity of SST retrievals to true changes in surface temperature should not be overlooked, especially in DV studies. The ideal SST retrieval will be unbiased, have low standard deviation of retrieval error, and will also have sensitivity to true SST variability close to 1.0. However, as shown in section 3, the overall sensitivity ratio of v3 MTSAT-1R data is 0.60 ± 0.05 , which is consistent with the findings in Merchant et al. (2009). Merchant et al. (2009, Figure 1c) indicates that Pathfinder Nonlinear SST (NLSST) retrievals from the AVHRR series have a sensitivity ratio of 0.4–0.7 over the TWP region. Considering this, and that the amount of available drifting buoy SST measurements is small over this study period and domain, this study has conducted all analyses based on both the original MTSAT-1R and the scaled estimated MTSAT-1R/ratio data. However, we recommend that the results based on the sensitivity-scaled MTSAT-1R/ratio data be considered more realistic.

The CG03 empirical model, which in this study is derived from the Centre de Météorologie (CMS) NLSST SEVIRI SST data, has the advantage that it is simple to use, requiring only wind speed and SSI as inputs. Compared with the results in Gentemann et al. (2003), over the TWP domain, this study's observed CG03 DV cycles and spatial distribution of DV signals are similar. Validated against the original MTSAT-1R data, the distribution shape of CG03 modeled $dSST_{max}$ best agrees with observations (Figure 5). Average CG03 DV cycles are also in reasonably close agreement with observations under most wind and DV situations, except for DV events > 2 K, which only occur with a 5.6% probability (Figure 5a). Spatially, CG03 captures the DV patterns in terms of both amplitude and distribution. However, when validated against the scaled MTSAT-1R data, severe underestimation is observed in CG03 for most DV and wind conditions, including the spatial distribution. One possible reason is that the CMS NLSST SEVIRI SST product, from which the CG03 model was calculated, has an overall sensitivity ratio of 0.72 (Merchant et al., 2013). Moreover, one notable deficiency for CG03 is that it estimates very few $dSST_{max} > 3$ K. This is possibly due to the model formulation which does not build on a previously calculated DV estimate. The CG03 model is based on instantaneous wind speed, and does not account for persistence; therefore, it will underestimate these strong events where persistent low winds result in exceptionally large DV values. As mentioned in section 2.3, another drawback of the SEVIRI CG03 is that it can only simulate $SST_{subskin}$ DV, due to the nature of its derivation.

When compared against the original MTSAT-1R data, ZB05 $dSST_{max}$ values are generally positively biased under MTSAT-1R $dSST_{max} < 2$ K conditions. Similar results are found in Bellenger and Duvel (2009) and Karagali and Høyer (2013) where the ZB05 modeled DV is compared with in situ and SEVIRI SST measurements, respectively. However, in the comparison with sensitivity-scaled MTSAT-1R/ratio data, ZB05 performs quite well in most of the months, under moderate DV (< 2 K) and wind speed conditions, and in terms of the

spatial distribution. This is largely because, in contrast to CG03, ZB05 is a physical model and the parameters are determined directly from in situ measurements. Soloviev and Lukas (1997) and Bellenger and Duvel (2009) argued that the temperature profile in the first 3 m used in ZB05 corresponds more closely to very strong diurnal warm layers. However, from the observed agreement between ZB05 and scaled MTSAT-1R data, it would appear that the ZB05 temperature profile is applicable to this study. When proposed, ZB05 has “realistically reproduced” a DV event > 2.5 K under low ECMWF wind conditions over a zonal band in the north-western Atlantic in May 1998 (Zeng & Beljaars, 2005). Our case study tells a similar story.

Takaya et al. (2010) refined the ZB05 model by changing the similarity function for strongly stable conditions and including the effect of the Langmuir circulation in calculating the turbulent diffusivity. In this study, when validated against the original MTSAT-1R data, ZB + T agrees closely with MTSAT-1R data in DV cycles under almost all conditions, except for the large negative biases for strong DV events, consistent with the warning in Takaya et al. (2010) of possible DV underestimation for very calm and clear sky conditions. However, serious underestimation is observed when comparing ZB + T with sensitivity-scaled MTSAT-1R/ratio data especially for calm winds ($< 3 \text{ m s}^{-1}$) and > 1 K DV conditions.

The air-sea coupled GC2 model is included in this study to evaluate the incorporation of DV schemes into a latest-generation coupled general circulation model, as the configuration used in this study implements the ZB05 warm layer scheme and a cool skin layer scheme based on Artale et al. (2002). Overall, GC2 has a large positive bias (0.72 K) when validated against the original MTSAT-1R data, and a much smaller positive bias (0.2 K) against the scaled MTSAT-1R/ratio data. Figure 5 indicates that, in both original MTSAT-1R and scaled MTSAT-1R/ratio cases, GC2 overestimates for most DV amplitudes, especially for DV < 2 K events. The frequent occurrence of DV events with < 2 K amplitude leads to GC2's overestimation for all months (Figure 6) and wind conditions (Figure 8). However, GC2 estimates DV > 2 K or 3 K better than the other models (Figures 7 and 10). Our case study shows that for extreme DV events, GC2 might be relatively accurate in terms of amplitudes. One issue in GC2 is the warming start (10–11 LST) and peak times (~ 17 LST), which are both 1–3 h later than the satellite measurements and other models. Further experiments indicate that this is not due to the lack of the cool skin scheme in the GC2 SSTs studied here, since including the cool skin effect does not change the warming start or peak times (figure not shown). Very late peak DV time leads to sharp temperature decline in the afternoon from 17 to 20 LST, followed by an unexpectedly constant temperature from 20 LST until midnight. The wind speed data comparison in section 4.1 indicates that GC2 winds are not responsible for the late DV start and peak times as GC2 winds are very similar to the ACCESS-R wind inputs for the other three DV models. The different behaviours in GC2, compared to the other models studied, must be due to other reasons.

In conclusion, if we accept the scaled MTSAT-1R/ratio data as most realistic, ZB05 may best capture the overall average dSST cycles and spatial distribution under most DV (maybe except for large DV events with amplitudes > 3 K) and most wind conditions. CG03 and ZB + T underestimate most DV events with amplitudes > 1 K. GC2, on the other hand, is positively biased for most DV events but able to predict reasonably well very large DV events, say > 3 K.

Appendix A: Derivation of V3 Geostationary MTSAT-1R SST

The MTSAT-1R satellite carried the Japanese Advanced Meteorological Imager (JAMI) and was in orbit above 140°E from 2005 through 2016. JAMI captured full-disc imagery on an hourly basis during the period 2005–2010 in five spectral channels (0.6–12.0 μm). As part of the Bureau's contribution to Australia's Integrated Marine Observing System (IMOS), observations from the JAMI spectral channels centred at 3.7, 10.8, and 12.0 μm were used to calculate SST using the NOAA-developed Geostationary Satellite Derived SST Processing System (Maturi et al., 2008). The NOAA software was modified to accept locally generated NWP fields and to estimate SST from the empirically derived relationship between JAMI observations and drifting buoy measurements of ocean temperature (Beggs et al., 2013). These IMOS version 2 (v2) MTSAT-1R SST estimates are available in $0.05^\circ \times 0.05^\circ$ gridded, hourly, single scene, GHRSSST format level 3 files for the period 1 June 2006 to 26 July 2010 from <http://rs-data1-mel.csiro.au/thredds/catalog/imos-srs/sst/ghrsst/L3U/mts1r/catalog.html>.

Initial comparison to the network of drifting buoys indicated that the difference between IMOS v2 MTSAT-1R and buoy SST observations varied spatially and temporally (Majewski et al., 2013; Figures A1a and A1b),

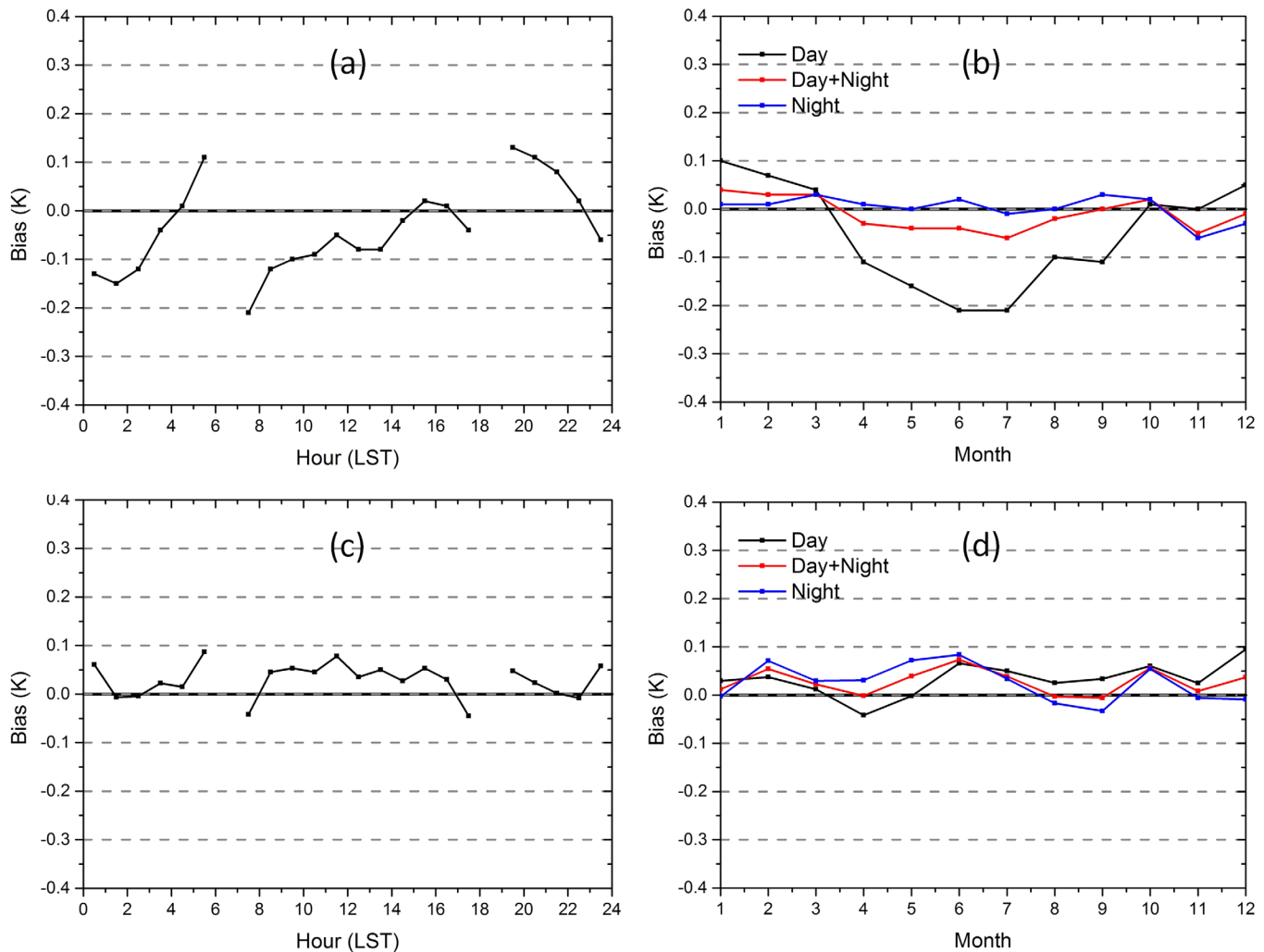


Figure A1. The median bias between IMOS MTSAT-1R and collocated drifting buoy SST observations as a function of (a, c) local hour, and (b, d) month, over the whole MTSAT-1R field of view. (top) Before applying the correction factors in equation (A1). (bottom) After applying the correction factors in equation (A1). *Note:* for the comparison, the buoy SST_{depth} measurements were converted to SST_{skin} by subtracting 0.17 K. The discontinuity in Figures A1a and A1c is due to the few good quality MTSAT-1R SST observations at dusk and dawn (see section 4.3).

with biases on the order of ± 0.2 K and standard deviations on the order of 0.8–1.2 K (Majewski et al., 2013). Additionally, the use of different algorithms for day (2-channel) and night (3-channel) scenes introduced hour-to-hour differences in the bias of >0.2 K. This order of uncertainty reduces the utility of the data for temporal studies of DV.

The spatial variation of the biases in the IMOS MTSAT-1R SSTs (Majewski et al., 2013) suggested that the issue might be due to the “fan-like” scanning pattern of the JAMI instrument caused by its two-axis gimbaled scan mirror system (Puschell et al., 2006). The observed temporal biases (Figures A1a and A1b) are hypothesized to be the result of solar contamination, depending on the solar declination and Earth-Sun distance. In order to reduce the temporal and spatial biases in the IMOS MTSAT-1R SST, the following correction factors were developed from a number of geometric and temporal properties, including pixel/line position, observation hour, solar declination and Earth-Sun distance:

$$\text{Corrected SST}_{\text{skin}} = \text{SST}_{\text{skin}} + \text{GFAC} + \text{DFAC} + \text{TFAC} \quad (\text{A1})$$

GFAC is a correction associated with the scan pattern (Puschell et al., 2006); DFAC is the SST correction caused by the solar declination/Earth-Sun distance; TFAC is a correction for the time of day, and these various SST components are parameterised as

$$SST_{skin} = p_0 T_4 + p_1 (T_4 - T_5) + p_2 (T_4 - T_5) \sec \theta \quad (A2)$$

$$GFAC = p_3 (XIDX - p_4)^2 + p_5 YIDX^2 \quad (A3)$$

$$DFAC = p_6 DECL + p_7 (ESDIST - 1) \quad (A4)$$

$$TFAC = p_8 \sin(\pi OBSHOUR/12) + p_9 \sin(2\pi OBSHOUR/12) \quad (A5)$$

where T_4 is the brightness temperature of channel 4 (11 μm), T_5 is the brightness temperature of channel 5 (12 μm), θ = satellite zenith angle, $XIDX$ = pixel number in longitude direction, $YIDX$ = pixel number in latitude direction, $DECL$ = solar declination, $ESDIST$ = distance between Earth and Sun, and $OBSHOUR$ = integer hour of observation in UTC. Parameters p_0 to p_9 were determined through cost minimization (Levenberg-Marquardt), where the cost term was composed of the difference between corrected MTSAT-1R SST and drifting buoy SST for the period 1 June 2006 to 26 July 2010 at a combination of time steps: instantaneous, hourly and monthly time scales. The matchups used in the minimisation were filtered by ACCESS-R wind speed (day: wind $> 6 \text{ m s}^{-1}$, night: wind $> 2 \text{ m s}^{-1}$) to minimize the observation of diurnal variation. Upon minimizing the cost function, the resulting values of p_0 to p_9 were found to be $p_0 = 1.00598$, $p_1 = 2.33557$, $p_2 = 0.504579$, $p_3 = 0.535103$, $p_4 = 421.957$, $p_5 = -0.313387$, $p_6 = -0.0093742$, $p_7 = 17.7333$, $p_8 = -0.166866$, $p_9 = 0.0570238$. The resulting maximum amplitude of the $DFAC$ term was $\pm 0.1 \text{ K}$ and the maximum amplitude of the $TFAC$ term was $\pm 0.2 \text{ K}$. The $TFAC$ fit was based on observation hour (not local time), so any effect from applying $TFAC$ will be smeared across the DV cycle. However, it is possible that applying the $TFAC$ correction term would reduce the size of the $dSST_{\text{max}}$ by around 0.2 K .

The application of these correction factors reduced the spatial and temporal differences between MTSAT-1R and drifting buoy SST observations (Figure A1). The resulting bias is $< 0.1 \text{ K}$ with a standard deviation of $\sim 0.7 \text{ K}$. Hour-to-hour differences in SST are also $< 0.1 \text{ K}$, with the exception of day/night transitions ($< 0.2 \text{ K}$). The corrected version 3 ("v3") MTSAT-1R SST_{skin} values were converted to the GHRSSST TWP+ data set level 3 format for the period 1 January to 30 April 2010, and are used in this study.

References

- Artale, V., Iudicone, D., Santoleri, R., Rupo, V., Marullo, S., & D'Ortenzio, F. (2002). Role of surface fluxes in ocean general circulation models using satellite sea surface temperature: Validation of and sensitivity to the forcing frequency of the Mediterranean thermohaline circulation. *Journal of Geophysical Research*, 107(C8), 3120. <https://doi.org/10.1029/2000JC000452>
- Atlas, R., Hoffman, R. N., Ardizzone, J., Leidner, S. M., Jusem, J. C., Smith, D. K., et al. (2011). A cross-calibrated, multiplatform ocean surface wind velocity product for meteorological and oceanographic applications. *Bulletin of the American Meteorological Society*, 92(2), 157–174. <https://doi.org/10.1175/2010BAMS2946.1>
- Beggs, H., Castro, S., Wick, G., Majewski, L., LeBorgne, P., Merchant, C., et al. (2012). Update on the GHRSSST Tropical Warm Pool Diurnal Variability (TWP+) project. In *Proceedings of the GHRSSST XIII Science Team meeting* (pp. 188–189). Reading, UK: GHRSSST International Project Office, University of Reading. Retrieved from <https://www.ghrssst.org/meetings/13th-international-ghrssst-science-team-meeting-ghrssst-xiii/>
- Beggs, H., Majewski, L., Griffin, C., Verein, R., Sakov, P., Huang, X., et al. (2013). Report to GHRSSST14 from Australia—Bluelink and IMOS. In *Proceedings of the GHRSSST XIV Science Team meeting* (pp. 104–121). Reading, UK: GHRSSST International Project Office, University of Reading. Retrieved from <https://www.ghrssst.org/meetings/14th-international-ghrssst-science-team-meeting-ghrssst-xiv/>
- Beggs, H., Zhong, A., Warren, G., Alves, O., Brassington, G., & Pugh, T. (2011). RAMSSA—An operational, high-resolution, regional Australian multi-sensor sea surface temperature analysis over the Australian region. *Australian Meteorological and Oceanographic Journal*, 61(1), 1–22.
- Bellenger, H., & Duvel, J.-P. (2009). An analysis of tropical ocean diurnal warm layers. *Journal of Climate*, 22(13), 3629–3646. <https://doi.org/10.1175/2008JCLI2598.1>
- Blockley, E., Martin, M., McLaren, A., Ryan, A., Waters, J., Lea, D., et al. (2014). Recent development of the Met Office operational ocean forecasting system: An overview and assessment of the new Global FOAM forecasts. *Geoscientific Model Development Discussions*, 7, 2613–2638. <https://doi.org/10.5194/gmd-7-2613-2014>
- Brunke, M. A., Zeng, X., Misra, V., & Beljaars, A. (2008). Integration of a prognostic sea surface skin temperature scheme into weather and climate models. *Journal of Geophysical Research*, 113, D21117. <https://doi.org/10.1029/2008JD010607>
- Clayson, C. A., & Bogdanoff, A. S. (2013). The effect of diurnal sea surface temperature warming on climatological air–sea fluxes. *Journal of Climate*, 26(8), 2546–2556. <https://doi.org/10.1175/JCLI-D-12-00062.1>
- Deschamps, P., & Frouin, R. (1984). Large diurnal heating of the sea surface observed by the HCMR experiment. *Journal of Physical Oceanography*, 14(1), 177–184. [https://doi.org/10.1175/1520-0485\(1984\)014<0177:LDHOTS>2.0.CO;2](https://doi.org/10.1175/1520-0485(1984)014<0177:LDHOTS>2.0.CO;2)
- Donlon, C., Casey, K., Gentemann, C., LeBorgne, P., Robinson, I., Reynolds, R., et al. (2009). *Successes and challenges for the modern sea surface temperature observing system*. Paper presented at the Proceedings of Ocean-Obs09: Sustained Ocean Observations and Information for Society (Vol. 2, Publ. WPP-306), Venice, Italy, September 21–25, 2009. <https://doi.org/10.5270/OceanObs09.cwp.24>
- Donlon, C., Minnett, P., Gentemann, C., Nightingale, T., Barton, I., Ward, B., et al. (2002). Toward improved validation of satellite sea surface skin temperature measurements for climate research. *Journal of Climate*, 15(4), 353–369. [https://doi.org/10.1175/1520-0442\(2002\)015<0353:TIVOS>2.0.CO;2](https://doi.org/10.1175/1520-0442(2002)015<0353:TIVOS>2.0.CO;2)
- Donlon, C., Robinson, I., Casey, K., & Vazquez-Cuervo, J. (2007). The global ocean data assimilation experiment high-resolution sea surface temperature pilot project. *Bulletin of the American Meteorological Society*, 88(8), 1197. <https://doi.org/10.1175/BAMS-88-8-1197>

Acknowledgments

This is a Sino-Australian Research Centre for Coastal Management Publication Number 43. H. Zhang has been financially supported by the China Scholarship Council since August 2014 for his PhD studies in Australia. The authors thank Sandra Castro, Gary Wick, and Andy Harris for their advice and encouragement throughout the TWP+ project. The authors acknowledge the very helpful feedback on the manuscript from Matthew Wheeler and Harry Hendon (Bureau of Meteorology). We also acknowledge the help from J. While, T. Johns, and M. Best in implementing the warm layer and cool skin scheme in the GC2 model. The authors thank the Bureau for providing the v3 MTSAT-1R (obtained from <http://dap.nci.org.au/thredds/remoteCatalogService?catalog=http://dapds00.nci.org.au/thredds/catalog/rr5/satellite/GHRSSST/MTSAT-1R/catalog.xml>), ACCESS-R (obtained from https://datacatalogue.nci.org.au/metadata/0401/f9907_4028_5233_6570), and RAMSSA (obtained from <https://protect-au.mimecast.com/s/p7qsCp8AXKsMmrV4SDk9wU?domain=podaac.jpl.nasa.gov>) data to the TWP+ Project. The iQuam drifting buoy data were obtained from <https://protect-au.mimecast.com/s/CR-RCWLJMvF9oWn1HnHvcy?domain=star.nesdis.noaa.gov>. The CCMP winds were obtained from NASA's Physical Oceanography Distributed Active Archive Centre (PO.DAAC; <https://podaac.jpl.nasa.gov/>). The ECMWF SSI data were obtained from <http://apps.ecmwf.int/datasets/data/interim-full-daily/levtype=sfc/>. The IMOS v2 MTSAT-1R data was sourced from the Integrated Marine Observing System (IMOS)—IMOS is a national collaborative research infrastructure, supported by the Australian Government. The data are available from <https://protect-au.mimecast.com/s/mLpICXlKNwF588Yzh6dgtX?domain=rs-data1-mel.csiro.au>. Finally, we thank the three anonymous reviewers whose comments have helped to make the paper clearer and more accurate. The research was funded, in part, by the ARC Centre of Excellence for Climate System Science (grant CE110001028).

- Fairall, C., Bradley, E. F., Godfrey, J., Wick, G., Edson, J. B., & Young, G. (1996). Cool-skin and warm-layer effects on sea surface temperature. *Journal of Geophysical Research*, 101(C1), 1295–1308. <https://doi.org/10.1029/95JC03190>
- Gentemann, C. L., Donlon, C. J., Stuart-Menteth, A., & Wentz, F. J. (2003). Diurnal signals in satellite sea surface temperature measurements. *Geophysical Research Letters*, 30(3), 1140. <https://doi.org/10.1029/2002GL016291>
- Gentemann, C. L., Minnett, P. J., & Ward, B. (2009). Profiles of ocean surface heating (POSH): A new model of upper ocean diurnal warming. *Journal of Geophysical Research*, 114, C07017. <https://doi.org/10.1029/2008JC004825>
- Halpern, D., & Reed, R. K. (1976). Heat budget of the upper ocean under light winds. *Journal of Physical Oceanography*, 6(6), 972–975. [https://doi.org/10.1175/1520-0485\(1976\)006<0972:HBOTUO>2.0.CO;2](https://doi.org/10.1175/1520-0485(1976)006<0972:HBOTUO>2.0.CO;2)
- Kantha, L. H., & Clayson, C. A. (1994). An improved mixed layer model for geophysical applications. *Journal of Geophysical Research*, 99(C12), 25235–25266. <https://doi.org/10.1029/94JC02257>
- Karagali, I., & Høyer, J. (2013). Observations and modeling of the diurnal SST cycle in the North and Baltic Seas. *Journal of Geophysical Research: Oceans*, 118, 4488–4503. <https://doi.org/10.1002/jgrc.20320>
- Karagali, I., & Høyer, J. (2014). Characterisation and quantification of regional diurnal SST cycles from SEVIRI. *Ocean Science*, 10, 745–758. <https://doi.org/10.5194/os-10-745-2014>
- Karagali, I., Høyer, J., & Hasager, C. (2012). SST diurnal variability in the North Sea and the Baltic Sea. *Remote Sensing of Environment*, 121, 159–170. <https://doi.org/10.1016/j.rse.2012.01.016>
- Karagali, I., Høyer, J. L., & Donlon, C. J. (2017). Using a 1-D model to reproduce the diurnal variability of SST. *Journal of Geophysical Research: Oceans*, 122, 2945–2959. <https://doi.org/10.1002/2016JC012542>
- Kawai, Y., & Wada, A. (2007). Diurnal sea surface temperature variation and its impact on the atmosphere and ocean: A review. *Journal of Oceanography*, 63(5), 721–744. <https://doi.org/10.1007/s10872-007-0063-0>
- Li, W., Yu, R., Liu, H., & Yu, Y. (2001). Impacts of diurnal cycle of SST on the intraseasonal variation of surface heat flux over the Western Pacific warm pool. *Advances in Atmospheric Sciences*, 18(5), 793–806.
- Majewski, L., Griffin, C., & Beggs, H. (2013). *Spatial and temporal biases in MTSAT-1R SST*. Paper presented as a poster at the 4th Asia Oceania Meteorological Satellite User's conference. Melbourne, Australia, October 9–11, 2013. <https://doi.org/10.13140/RG.2.1.3861.6805>
- Marullo, S., Minnett, P., Santoleri, R., & Tonani, M. (2016). The diurnal cycle of sea-surface temperature and estimation of the heat budget of the Mediterranean Sea. *Journal of Geophysical Research: Oceans*, 121, 8351–8367. <https://doi.org/10.1002/2016JC012192>
- Masson, S., Terray, P., Madec, G., Luo, J.-J., Yamagata, T., & Takahashi, K. (2012). Impact of intra-daily SST variability on ENSO characteristics in a coupled model. *Climate Dynamics*, 39(3–4), 681–707. <https://doi.org/10.1007/s00382-011-1247-2>
- Maturi, E., Harris, A., Mittaz, J., Merchant, C., Potash, B., Meng, W., et al. (2008). NOAA's sea surface temperature products from operational geostationary satellites. *Bulletin of the American Meteorological Society*, 89(12), 1877–1888. <https://doi.org/10.1175/2008BAMS2528.1>
- Megann, A., Storkey, D., Aksenov, Y., Alderson, S., Calvert, D., Graham, T., et al. (2014). GO5.0: The joint NERC-Met Office NEMO global ocean model for use in coupled and forced applications. *Geoscientific Model Development*, 7(3), 1069–1092. <https://doi.org/10.5194/gmd-7-1069-2014>
- Merchant, C., Harris, A., Roquet, H., & Le Borgne, P. (2009). Retrieval characteristics of non-linear sea surface temperature from the Advanced Very High Resolution Radiometer. *Geophysical Research Letters*, 36, L17604. <https://doi.org/10.1029/2009GL039843>
- Merchant, C., Le Borgne, P., Roquet, H., & Legendre, G. (2013). Extended optimal estimation techniques for sea surface temperature from the Spinning Enhanced Visible and Infra-Red Imager (SEVIRI). *Remote Sensing of Environment*, 131, 287–297. <https://doi.org/10.1016/j.rse.2012.12.019>
- Noh, Y., Lee, E., Kim, D. H., Hong, S. Y., Kim, M. J., & Ou, M. L. (2011). Prediction of the diurnal warming of sea surface temperature using an atmosphere-ocean mixed layer coupled model. *Journal of Geophysical Research*, 116, C11023. <https://doi.org/10.1029/2011JC006970>
- Petrenko, B., Ignatov, A., Kihai, Y., Stroup, J., & Dash, P. (2014). Evaluation and selection of SST regression algorithms for JPSS VIIRS. *Journal of Geophysical Research: Atmospheres*, 119, 4580–4599. <https://doi.org/10.1002/2013JD020637>
- Price, J. F., Weller, R. A., Bowers, C. M., & Briscoe, M. G. (1987). Diurnal response of sea surface temperature observed at the long-term upper ocean study (34°N, 70°W) in the Sargasso Sea. *Journal of Geophysical Research*, 92(C13), 14480–14490. <https://doi.org/10.1029/JC092iC13p14480>
- Price, J. F., Weller, R. A., & Pinkel, R. (1986). Diurnal cycling: Observations and models of the upper ocean response to diurnal heating, cooling, and wind mixing. *Journal of Geophysical Research*, 91(C7), 8411–8427. <https://doi.org/10.1029/JC091iC07p08411>
- Puri, K., Dietachmayer, G., Steinle, P., Dix, M., Rikus, L., Logan, L., et al. (2013). Implementation of the initial ACCESS numerical weather prediction system. *Australian Meteorological and Oceanographic Journal*, 63, 265–284.
- Puschell, J. J., Osgood, R., Auchter, J., Hurt, W. T., Hitomi, M., Sasaki, M., et al. (2006). *In-flight performance of the Japanese Advanced Meteorological Imager*. Paper presented at the Proceedings of the SPIE 6296, Earth Observing Systems XI, 62960N. <https://doi.org/10.1117/12.683505>
- Soloviev, A., & Lukas, R. (1997). Observation of large diurnal warming events in the near-surface layer of the western equatorial Pacific warm pool. *Deep Sea Research Part I: Oceanographic Research Papers*, 44(6), 1055–1076. [https://doi.org/10.1016/S0967-0637\(96\)00124-0](https://doi.org/10.1016/S0967-0637(96)00124-0)
- Soloviev, A., & Lukas, R. (2013). *The near-surface layer of the ocean: Structure, dynamics and applications* (Vol. 48). Berlin, Germany: Springer Science & Business Media.
- Stuart-Menteth, A. C., Robinson, I. S., & Challenor, P. G. (2003). A global study of diurnal warming using satellite-derived sea surface temperature. *Journal of Geophysical Research*, 108(C5), 3155. <https://doi.org/10.1029/2002JC001534>
- Takaya, Y., Bidlot, J. R., Beljaars, A., & Janssen, P. A. (2010). Refinements to a prognostic scheme of skin sea surface temperature. *Journal of Geophysical Research*, 115, C06009. <https://doi.org/10.1029/2009JC005985>
- Wang, X. H., & Zhang, H. (2017). Effects of Australian summer monsoon on sea surface temperature diurnal variation over the Australian North-Western Shelf. *Geophysical Research Letters*, 44, 9856–9864. <https://doi.org/10.1002/2017GL075008>
- Webster, P. J., Clayson, C. A., & Curry, J. A. (1996). Clouds, radiation, and the diurnal cycle of sea surface temperature in the tropical western Pacific. *Journal of Climate*, 9(8), 1712–1730. [https://doi.org/10.1175/1520-0442\(1996\)009<1712:CRATDC>2.0.CO;2](https://doi.org/10.1175/1520-0442(1996)009<1712:CRATDC>2.0.CO;2)
- Weih, R. R., & Bourassa, M. (2014). Modeled diurnally varying sea surface temperatures and their influence on surface heat fluxes. *Journal of Geophysical Research: Oceans*, 119, 4101–4123. <https://doi.org/10.1002/2013JC009489>
- Williams, K., Harris, C., Bodas-Salcedo, A., Camp, J., Comer, R., Copsey, D., et al. (2015). The met office global coupled model 2.0 (GC2) configuration. *Geoscientific Model Development*, 8(5), 1509–1524. <https://doi.org/10.5194/gmd-8-1509-2015>
- Xu, F., & Ignatov, A. (2014). In situ SST quality monitor (iQuam). *Journal of Atmospheric and Oceanic Technology*, 31(1), 164–180. <https://doi.org/10.1175/JTECH-D-13-00121.1>

- Zeng, X., & Beljaars, A. (2005). A prognostic scheme of sea surface skin temperature for modeling and data assimilation. *Geophysical Research Letters*, 32, L14605. <https://doi.org/10.1029/2005GL023030>
- Zhang, H., Beggs, H., Majewski, L., Wang, X. H., & Kiss, A. (2016a). Investigating sea surface temperature diurnal variation over the Tropical Warm Pool using MTSAT-1R data. *Remote Sensing of Environment*, 183, 1–12. <https://doi.org/10.1016/j.rse.2016.05.002>
- Zhang, H., Beggs, H., Wang, X. H., Kiss, A. E., & Griffin, C. (2016b). Seasonal patterns of SST diurnal variation over the Tropical Warm Pool region. *Journal of Geophysical Research: Oceans*, 121, 8077–8094. <https://doi.org/10.1002/2016JC012210>